INSIDE *AND* OUTSIDE THE BLACK BOX

The Constraining Effect of an Object Boundary on Illusory Conjunctions

BRONWYN H. RUTHERFORD

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Psychology.

UNIVERSITY OF CANTERBURY, 1992.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Dr. Gill Rhodes, for her support and encouragement, as well as for suggestions and help in clarifying my own ideas. I would also like to offer her special thanks for her continued supervision and support from the U.S.A. (via FAX machine and electronic mail - not an easy task!) after going overseas on sabbatical.

I would also like to specially thank Mr. Paul Russell for his valued help and suggestions, particularly with regard to visual attention, in Gill's absence. His kind support and the time taken to read drafts of this thesis, discuss its contents, and make suggestions at exam marking time was very much appreciated.

Finally, I would like to thank my family: my daughters, Cara, Renee (and her partner Mark), Keryn, and Nina, for their tolerance and faith in me. Also my ex-partner, Ian Hutchens, for his encouragement and financial help.

CONTENTS

ABSTRACT

Illusory conjunctions were originally offered as part of the convergent evidence that supported Feature Integration Theory (FIT) (Treisman, 1988; Treisman & Schmidt, 1982). According to FIT, illusory conjunctions arise from a lack of focal attention. An extended version of Prinzmetal and Keysar's (1989) functional explanation of illusory conjunctions is offered as an alternative to FIT's account of illusory conjunctions. This functional explanation rests on the hypothesis that illusory conjunctions are the result of an adaptive response to very brief viewing times in the presence of visual location information that constrains both feature integration and visual attention. Two experiments were conducted to test predictions arising from this hypothesis. In these experiments the effects of a black square outline on illusory conjunction and non conjunction error rates were examined. The results of the experiments indicate that the critical factor in eliciting illusory conjunctions is the presence of constraining visual information. It is also demonstrated that such constraints not only lead to illusory conjunctions, but also facilitate the accurate detection of objects. The apparent conflict in the claim that constraints lead to illusory conjunctions, which are errors, and at the same time facilitate accuracy can be resolved by the functional explanation. Some suggestions are made as to which visual stimuli might operate as constraints on feature integration and visual attention.

1
INTRODUCTION

Treisman and Schmidt (1982) used the term *illusory conjunctions* to describe errors made in visual cognition, that combine features from separate objects present in a visual scene, so they are perceived as a whole coherent object that is not actually present. This thesis is primarily concerned with the phenomenon of illusory conjunctions and explanations of why they occur. However, first an attempt will be made to put the research related to this phenomenon into historical and theoretical context.

Visual experience starts with retinal images. Our subjective visual experiences are, however, vastly different to the two-dimensional, upside-down images that fall on the retina. Theories of visual cognition have attempted to explain how these retinal images are transformed into the complex three-dimensional, often moving, visual images we experience. Template Matching models were early attempts to explain how we perceive and identify objects and events in the world. These models proposed that we directly compare retinal images to various stored patterns (or templates) in order to organize our visual perceptions (Anderson, 1985). However, human visual experience is extremely complex and flexible, and the possible number of visual patterns that could be stored is probably infinite. Template Matching models cannot account for this flexibility (Anderson, 1985), and cognitive economy makes it unlikely that a finite brain would deal with stimuli in such a manner (Treisman, 1985).

Other theories are more complex and have postulated different stages, or levels, of perception. There are two main groups of these more complex theories. First, the Gestalt theories propose that our initial perception of a scene or object is whole, and that this whole unit is analyzed "down" into its parts when necessary. They further propose that certain innate principles organize stimuli into whole units and determine how we segment an object into its components. These principles include proximity, similarity, good continuation, closure, and good form (Anderson, 1985). There is evidence to support this view, including evidence that suggests we may perceive larger configurations faster and more accurately than their components. For example, Pomerantz, Sager and Stoever (1977) showed that subjects recognized configurations faster than their parts.

In contrast to the Gestalt theories, the Feature Analysis models propose that perceived stimuli are combinations of basic elements such as lines, angles and curves, and that objects are identified by analyzing these elements, or features, into recognizable patterns (Anderson, 1985; Lindsay & Norman, 1872). In effect, stimuli are initially encoded as a set of separate features which are then combined to form a whole unit or object (Treisman, 1985). Hubel and Wiesel (1962) discovered that individual cortical cells in the visual cortex of cats were activated only by lines or edges of a particular orientation and width, offering considerable support for the Feature Analysis models of visual cognition. However, while the Gestalt theories provide a set of principles that guide how we might analyze initially perceived whole units into their component parts, early Feature Analysis models did not specify how initially detected features might be combined to form the whole objects we subjectively experience in everyday life. In 1980 Treisman and Gelade produced an attentional theory of feature integration that did specify how features might be combined to form objects, and postulated that attention plays a necessary role in this cognitive processing. Treisman and her colleagues have provided convergent evidence to support their Feature Integration Theory (FIT), including research related to visual search, texture segregation, feature location, and illusory conjunctions. Before taking a detailed look at the research and explanations related to illusory conjunctions, and how visual attention might be related to this phenomenon, Treisman's FIT and the evidence to support it will be described. Some alternatives to FIT will also be discussed in the section dealing with feature integration. Because FIT holds attention necessary for combining features into objects, and there is considerable debate with regard to this hypothesis, the first section will be assigned to a brief discussion of the nature of visual attention.

2
VISUAL ATTENTION

This section is relatively brief and will not describe all of the evidence related to visual attention issues. Although visual attention is an important feature of the visual cognition theories being discussed, it is not the central issue, and to cover it fully would require a far bigger review of the literature than is within the scope of this thesis.

Attention has been conceptualized as a limited resource, and studies on sensory memory have indicated that although a large amount of information gets into sensory memory, it is quickly lost if not attended to. The results of research, such as dichotic listening experiments, have been interpreted as indicating that attention has an important role in selecting sensory information for further processing (Anderson, 1985). FIT proposes that attention is necessary for combining features into whole objects but this claim is controversial (e.g. Allport, 1989; Prinzmetal & Keysar, 1989; Tsal, 1989). However, two main issues, both of which FIT makes important assumptions about, need to be addressed before considering how visual attention and visual processing might be related. The first issue is how visual attention is moved, spread, or allocated across a visual scene. The second is related to what captures attention, and whether visual attention is spatially or object based.

HOW VISUAL ATTENTION IS ALLOCATED

The nature of attention has been characterized in two different ways (Humphreys & Bruce, 1989). First, as a "constant velocity spotlight" (Cheal & Lyon, 1989), and more recently as something that operates more like a "zoom lens" (Treisman and Gormican, 1988). The spotlight model of attention proposed that visual attention was fixed in size or spread, could not be split or divided between non-adjacent locations, and was moved serially at a constant rate across the visual field (Treisman, 1985). There was evidence to support this notion (Erikson & Yeh, 1985; Shulman, Remington & McLean, 1979; Treisman & Gelade, 1980; Tsal 1983) but subsequent research has shown that the second account may be more accurate. This more recent account proposes that attention operates along a gradient (Downing & Pinker, 1985; LaBerge & Brown, 1986; LaBerge & Brown, 1989) more like a "zoom lens" that may be spread over a whole scene, part of a scene, or narrowed down to focus on a small area, thereby improving resolution (Cheal & Lyon, 1989; Humphreys & Bruce, 1989; Murphy & Erikson, 1987). It should be noted that attentional focus is differentiated from eye fixation, as it is possible to attend to visual input from peripheral vision, as well as to visual stimuli that fall on or near the foveal fixation point (Posner, 1980).

Both the spotlight and zoom lens accounts of visual attention seem to imply that once attention is fixed on a certain area of space all the stimuli within that area will be processed while stimuli outside of the area will not. Treisman (1988) has questioned this notion, and there is evidence to suggest that not all stimuli within an attended space are processed equally, but that some information may be "filtered out" and other information may be processed together even when parts of it may be separated by unattended, or filtered out, items. (Nakayama & Silverman, 1986; Prinzmetal & Keysar, 1989; Treisman, 1988). LaBerge and Brown (1989) have proposed a Gradient Model of attention that includes a filtering process as well as an hypothesis regarding how attention might be shifted from one area of visual space to another. They suggest that in visual search experiments (in which subjects are required to detect a target among varying numbers of distractor items) top-down expectancies will lead to a filtering mechanism inhibiting the input from items adjacent to a target once attention is fixed on the expected target location. So in their model LaBerge and Brown suggest a mechanism by which some items may be "filtered out" and other items processed, even when they are all fall under the attentional "lens." The theory proposes that location and feature information are encoded separately. The filter operates on the location information and determines which features in a display are the ones to be identified.

Initial conceptions of the zoom lens model of attention also assumed that attention was moved serially across the visual scene in an analogue of real motion, when it was shifted from one location to another, in the same way it was assumed for the spotlight model (Downing & Pinker, 1985; LaBerge & Brown, 1986). However, LaBerge and Brown (1989) propose that the attentional gradient is not moved in this way but rather, that a new gradient is formed each time a new location is attended to. There is evidence to support this proposal (LaBerge & Brown, 1989), including clear evidence that attention does not move at a constant velocity across a visual scene (Cheal & Lyon, 1989; Egly & Homa, 1991).

Attentional boundaries.

Both the spotlight and zoom lens accounts of attention spread suggest that attention may have clear boundaries (Cohen & Ivry, 1989; Treisman, 1988). That is, when attention is focused on some object or area of space, other areas of space will fall outside of attention, implying an attentional boundary exists. One possibility is that attentional boundaries are determined by form outlines, such as the box cue used in selective cueing experiments like those of Nakayama and Mackeben (1989). Another possibility is that there are no clear or sharp boundaries to attention, but rather that such boundaries may be fuzzy (Cohen & Ivry, 1989) and flexible (LaBerge & Brown, 1989). LaBerge and Brown, in their Gradient Model of visual attention, propose that attention does not have clear boundaries and their research, as well as others', indicates that there is a gradual decrease in the quality of visual detection of items as they are presented further away from an attention cue (e.g. Downing, 1988).

ATTENTION CAPTURE

The second issue in visual attention that is important for visual cognition theories, is attention capture. Attention capture incorporates two related issues: whether visual attention is spatially or object based, and what its adaptive function is. Whether visual attention is spatially or object based has been related to whether attentional selection is "early" or "late" in visual processing (Allport, 1989). That is, whether it occurs at a relatively early stage of visual analysis prior to object recognition (Treisman, 1988; Prinzmetal & Keysar, 1989), or whether it occurs later in the visual process at the semantic, conceptual, or object representation levels of processing (Allport, 1989; Duncan, 1980; Shiffrin & Schneider, 1977). If attentional selection occurred very early in visual processing, perhaps before features have been combined into objects, then it is reasoned that attention must be spatially based, because it could not be captured by objects if the visual system had not yet registered the visual information it was receiving as whole coherent objects. Further, attention would be captured by location information, once again because other visual information would not have been processed enough to have transformed the features, or basic elements, of objects into the whole entities we experience in human vision (Treisman & Gelade, 1980). The other view is that attention would not select visual items until visual information was processed enough for objects to be recognized, or have meaning (Allport, 1989). The debate over whether attentional selection occurs early or late in visual processing has not been resolved (Allport, 1989), but this question is likely to be related to what the function of attention is, or for what purpose in evolutionary adaption it was selected.

Allport (1989) proposes that, rather than characterizing attention as a limited capacity resource that supposedly "enables certain sorts of processing" (p662), it might be better conceived of as a multimodal function that selects for action, or selects "for the potential control of action" (p649). If attention is a "selection-for-action" one might surmise that attention is object based because organisms act on, or interact with, their environment. Moreover, they interact with objects or separate areas of space (e.g. a lawn) in their environment. As already mentioned, the controversy over whether attentional selection occurs early or late in visual processing has not been empirically resolved. Allport suggests that a more important question regarding attention capture could be what mechanisms, both internal and external, might serve to engage (or capture) and constrain visual attention. It is possible that in learning what these mechanisms are, the controversy over whether attentional selection occurs early or late in visual processing might be resolved. Research in visual attention has found evidence that some stimuli are more likely to capture attention than others (Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990) and that there may be more than one type of attention.

Posner (1980) makes a distinction between two systems that might serve in the capture of attention: an involuntary, externally manipulated, exogenous attentional system and a voluntary, internally controlled, endogenous attentional system. It is likely that these attentional systems would interact (Humphreys & Bruce, 1989). A set of experiments conducted by Nakayama and Mackeben (1989) provide evidence for two attentional mechanisms; a transient involuntary component and a sustained controlled component. In these experiments subjects were presented with horizontal and vertical black or white bars, and required to detect the colour of an odd target if one was present. In some instances the target was uncued, while in others it was cued by a box that surrounded the target area. Cues were concurrent and sustained on some trials, but were transient precues on other trials (i.e. the cue was presented before the target but was not present during the target display). Target detection was enhanced by cueing, but the amount of performance improvement differed for the two cued conditions. Precueing enhanced performance more than sustained concurrent cueing, but only if it occurred between 70ms and 150ms before the target display onset. Longer delays than this between cueing and display onset led to a downturn in performance, indicating a transient attentional component. However, the result of interest is that performance was facilitated more by a transient precue than a sustained cue. It is likely that internally controlled top-down processes, resulting from prior knowledge of the target, would lead to the selection of a target, but it is also apparent that the box cue in this experiment had some automatic effect on the attentional selection of the target area, before the target had been presented when the cue was a transient precue. It is also interesting to note that when a cue was sustained through-out the target presentation, performance was not as good as when the cue was a transient pre-cue. Downing (1988) suggests that the box cues used in selective cuing experiments may contribute to the results by interfering with performance. It is possible that this down-turn in performance with the sustained cue could, at least in part, be explained by the cue interfering with target detection when it remained present throughout the display time. That is, the box cue may not only automatically capture attention, but may also mask the target to some degree by automatically taking priority in processing over the target if it remains present during the target display.

Although a briefly presented peripheral stimulus seems to engage attention automatically, and interfere with the detection of a centrally displayed target (Posner, Snyder, & Davidson, 1980), suggesting peripheral movement or sudden onset engages the exogenous system, it is also likely that characteristics of a centrally fixated visual scene would also engage this exogenous system. Novel stimuli have been shown to capture attention over other more familiar stimuli in a display (Johnston et al., 1990), but as Johnston et al. observe, little is known about what might capture attention when no novel stimuli are present (as in a familiar scene) or when the whole scene is unfamiliar. Given Posner's (1980) suggestion and Nakayama and Mackeben's evidence that attention is controlled both by external and internal factors, it is probable attentional selection occurs as a result of top-down processes. This notion is accounted for in LaBerge and Brown's (1989) Gradient Model, in their suggestion that prior knowledge of a target will lead to inhibition of activation from items adjacent to the target. However, it is also likely that external factors would control attention, not only in the absence of top-down processes, but even when there is prior knowledge or expectancies guiding attention; a notion supported by Nakayama and Mackeben's results. Allport (1989) suggests it is possible that Gestalt organizational principles, like grouping, proximity and closure, could be involved in early visual selection in the absence of novel stimuli; a notion which is also in keeping with Nakayama and Mackeben's (1989) results in which a box cue around the target not only captured attention, but also continued to have some effect on perception of the attended area if the cue was sustained throughout the display time. Allport's proposal also suggests that external factors would actually constrain attention even when top-down processes, such as intentions, are involved in attentional selection. That is, something like the Gestalt grouping principles may cause attention to be applied to data from a visual scene only in certain ways. For example, even when intentions lead to internal, top-down, controls on which item in a visual field is selected, this internally controlled selection would also be partly controlled (or constrained) by external factors: It would be more adaptive to attend to particular objects, or areas of space like a lawn, rather than the spaces between objects or the collected parts of various objects. As Allport points out, organisms interact with objects, so it would be adaptive to have attention select objects rather than just any aspect of visual space, whether selection occurs late or early in visual processing. It also follows that there would be cues from the environment that are important in separating objects from space as well as from each other.

Allport's (1989) functional view of attention as "selection-for-action", as well as the idea that attention is constrained by cues related to environmental factors, is supported by a phenomenon, called "motion capture" by Ramachandran (1990), and which he observed in some of his experiments. When subjects in these experiments were shown apparently moving gratings with a background of stationary random dots, the dots were perceived as moving in the same direction as the gratings. Even illusory squares, formed by four incomplete disks, captured a matrix of dots enclosed within their boundaries into their apparent motion. Ramachandran offers a functional explanation of his results. He suggests that in perceiving something like a leopard, the visual system extracts motion signals from conspicuous properties, like the form outline, and discards (or filters out) information about functionally less relevant details, like the leopard's spots. He points out that it would be more adaptive to perceive a leopard's form and movement rather than details like its pattern of spots. Rhamachandran's functional explanation, like Allport's, and the results of his experiments, not only suggest that external factors might capture and constrain attention, they also suggest that visual cognition may not treat all input from a single object or area equally, even when that input all falls under the attentional "spotlight" or "lens."

SUMMARY

Current evidence suggests that visual attention is not a constant velocity spotlight with clear or sharp boundaries, that is moved serially across the visual scene in an analogue of real motion. The evidence indicates that it is a gradient that operates more like a zoom lens, has fuzzy or unclear boundaries, and is relocated by the formation of a new gradient at each attended location, rather than being moved across a visual scene (LaBerge & Brown, 1989).

Current evidence also indicates that there are two types of attention: an involuntary, externally manipulated, exogenous mechanism and an internally controlled, endogenous mechanism. It is apparent that both peripherally and centrally located novel stimuli engage the exogenous attentional system, thereby capturing attention, but it is not clear what might capture attention in the absence of such novel stimuli

(Johnston et al., 1990). Also, although the debate over whether attention is object or spatially based is as yet unresolved, it is possible that identifying what constrains visual attention may help to clarify this issue (Allport, 1989). It is possible that cues related to environmental factors (like the Gestalt organizational principles) could be involved in early visual selection, by constituting at least some of the constraints on visual attention, and that these constraints would operate whether attentional selection has resulted from automatic, externally controlled processes, or from internally controlled (i.e. top-down) processes.

In the next section feature integration and FIT will be discussed. As FIT holds visual attention to be necessary to feature integration, the issues of the movement and spread of attention, attention capture, and early versus late selection, are important issues about which FIT makes certain assumptions. These assumptions about visual attention will be pointed out as the theory is described.

3

FEATURE INTEGRATION

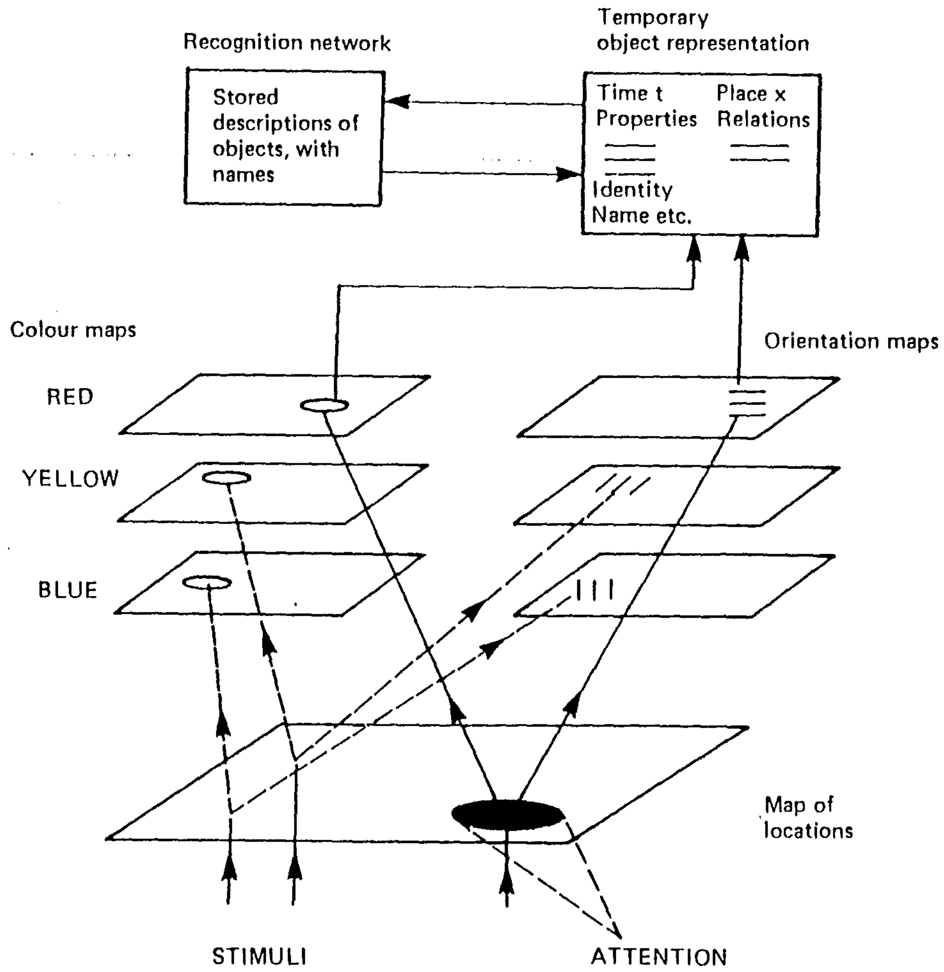
This section will centre on a description and discussion of Treisman's FIT. Evidence to support the theory, as well as problems posed for FIT by some related research, will be included. Some alternative theories and explanations will also be discussed, including Julesz' Texton Theory (Julesz, 1975, 1981, 1984) and Wolfe, Cave and Franzel's (1989) Guided Search Model.

FEATURE INTEGRATION THEORY

Treisman's FIT proposes that the visual scene is initially coded along a number of separate dimensions such as colour, orientation, brightness, and direction of movement, and that features or values along these dimensions, such as red or vertical, are coded separately in specialized modules or maps. Treisman and Gelade (1980) also suggest that relations between features could lead to "emergent features" that might be coded as primitive features by the visual system: that is, feature detectors might be "hard wired" in for some relations between features in the same way as for simple features like orientation or colour. It is postulated that features are coded automatically, and spatially in parallel, in a preattentive stage. It is also postulated that in the early stage of processing a "master map of locations" (Treisman & Gormican, 1988, p17) encodes *where* all the features are located but not *which* features are located where. So in the first stage of visual processing, empty locations are distinguished from filled locations, and feature representations and location information are encoded separately. In an elaboration of FIT, Treisman and Gormican (1988) proposed that some location information would be available preattentively. They suggest that features along a single dimension may be organized, within their own spatial map, in terms of how they are located in relation to each other. For example, the colour map would encode the locations of several colours in terms of how those colours related to each other spatially.

The theory claims that the location of features (from the different dimensions), and therefore how these features are conjoined, is only made available to conscious experience through their links to the areas in the master map that currently fall under the attentional spotlight. Attention is, therefore, *assumed to be spatially based* (i.e. the early selection model of attention is assumed) and also necessary as the "glue" that conjoins initially encoded feature representations into unitary object representations. That is, once an area of visual space is attended to, the separate features abstracted by the specialized feature maps in the preattentive stage, are recombined to form the objects and events that we subjectively experience. The final stage in the visual processes postulated by Treisman's FIT is described as one in which the conscious perception of objects results when these temporary object representations are matched to stored descriptions in long-term visual memory (Treisman, 1985; Treisman 1988; Treisman & Gelade, 1980; Treisman & Gormican, 1988). A copy of a diagram of the model of processing described in FIT is provided in Figure 1.

Figure 1: A diagram of the model of visual processing proposed by Treisman and her colleagues (copied from Treisman, 1988, p202).



In their initial formulation of FIT, Treisman and Gelade conceived of attention in terms of a dichotomy; that is, either attention or non-attention. They also assumed a constant velocity "spotlight" model of attention, proposing that attention is moved from one location to another by moving it so it passes over intermediate locations in an analogue of real motion. Subsequent evidence led Treisman and Gormican (1988) to surmise that attention probably operates along a continuum from very narrowly and finely focused to widely and diffusely spread: that is, the "zoom lens" account of visual attention discussed in the section on visual attention. They also assumed that the attentional lens has clear or sharp boundaries.

Evidence to support Treisman's FIT is convergent and related to two sets of predictions: those concerned with the detection of single features, and those concerned with the detection of items or objects that constitute a conjoining of features (i.e. feature conjunctions). The theory predicts that the detection of features should occur in parallel and without their necessarily being located. Features should also mediate easy texture segregation, and could be combined to form illusory conjunctions in the absence of, or when there is insufficient, attention. Conversely, detecting feature conjunctions should require focused attention, and that the items be located. Their detection should, therefore, require a serial search of filled locations in the visual field. Feature conjunctions would not be expected to mediate texture segregation (Treisman & Gelade, 1980). In the following four sub-sections, evidence from the four experimental paradigms used to test these predictions will be described, including visual search, feature location, texture segregation and illusory conjunction experiments. Some problems posed for FIT by evidence from other researchers will be discussed, as well as elaborations of FIT, where relevant.

VISUAL SEARCH

Treisman and Gelade (1980) used a visual search task to test the prediction that although feature detection might occur in parallel without focused attention, the detection of feature conjunctions would require serial search and focused attention. In this task subjects were asked to find randomly placed coloured letter targets in displays of distractors, which were also coloured letters but differed from the target letters. Stimuli were displays of brown T and green X distractors, with a target present in some displays, and were presented by tachistoscope. On positive trials the displays contained a target letter, while on negative trials they did not. Subjects indicated when a target was present by pushing an appropriate key. There were two target conditions: a feature condition in which the target was a blue T or X, or a brown or green S, so the target differed from both distractors by one feature; and a conjunction condition in which the target was a green T which contains a feature from each distractor. Displays varied over four sizes, containing 1, 5, 15, or 30 items. Search times were compared for the two conditions, and for positive and negative trials within those conditions. Figure 2 shows the distractors, feature targets and conjunction target more clearly.

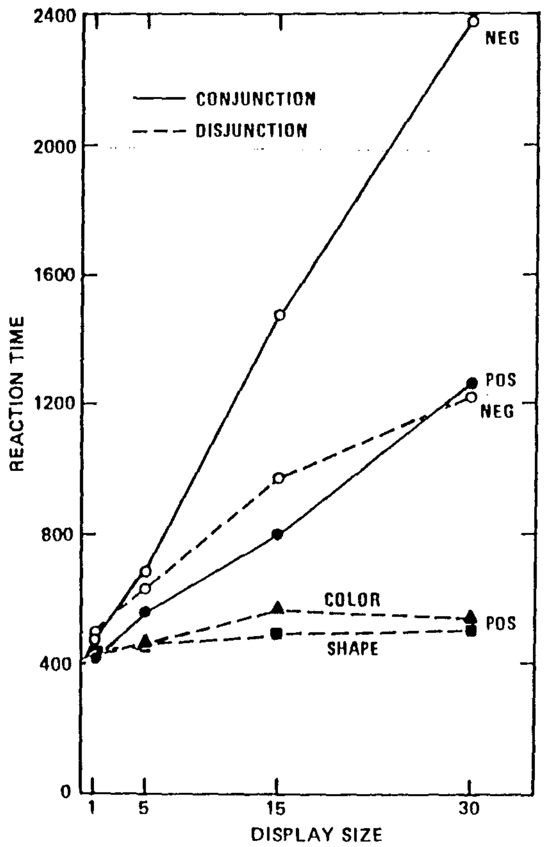
Figure 2: The distractors, feature targets and conjunction target utilized in Treisman and Gelade's (1980) visual search experiments.

Distractors:	T$^{\text{brown}}$	X$^{\text{green}}$
Feature targets:	T$^{\text{blue}}$	X$^{\text{blue}}$
	S$^{\text{brown}}$	S$^{\text{green}}$
Conjunction target:	T$^{\text{green}}$	

Treisman and Gelade's results indicated that in the conjunction condition search times increased linearly with display size, and for negative trials this increase was approximately double that for positive trials. The authors point out that if finding a target required serial search then one would expect search times to increase linearly with display size, and that negative trials would take longer as searches would be exhaustive rather than terminated by finding the target. If features can be detected in parallel then one would expect display size to have little effect on search times and, as predicted, this was the case in the feature condition. Search times for feature targets remained relatively constant across all display sizes, producing almost flat functions; a phenomenon termed "pop-out" (Treisman, 1988; Treisman & Gormican, 1988). Figure 3 (page 21) shows the mean search times for subjects in each condition in Treisman and Gelade's first experiment.

In another experiment Treisman and Gelade tested the effects of feature discriminability in conjunction search. They did this to show that serial search for targets is not attributable to tasks merely being more difficult, but rather to some qualitative difference in the way features and conjunctions are processed. They used two sets of stimuli that differed in discriminability in both shape (O and N versus T and X) and colour (red and green versus blue and green). They found that although identifying less discriminable targets took longer, the 2:1 ratio in search times (e.i. the qualitative difference) for positive and negative trials was preserved.

Figure 3: Mean search times as a function of display size for conjunction and feature (disjunction) search on positive (target present) and negative (target absent) trials (copied from Treisman & Gelade, 1980, p104).



In yet another experiment Treisman and Gelade explored the notion that the difference in times taken to identify feature and conjunction targets could be explained by the greater similarity the conjunction targets had to the distractors; a factor seen as critical by some researchers (e.g. Duncan & Humphreys, 1989). In this experiment Treisman and Gelade presented very large, very small or medium sized target ellipsis in distractor displays of moderately small and moderately large ellipses. They reasoned that the very large and very small ellipse targets should be more easily discriminated than the medium sized ellipse target, as the medium sized one was more similar to both distractors. In fact this was not the case, and also, the pattern of results was much different than for either the feature or colour-shape conjunction conditions in their previous experiments. The functions relating reaction time to display size were negatively accelerated, rather than the linear functions obtained for the feature and conjunction letter targets. The authors concluded that although search for similar targets might be slow, the fact that the function was not linear shows that search to discriminate among similar targets is not serial. However, although this experiment indicates that the difference in search latencies for conjunction and feature targets is probably not attributable to similarity, the slower search times indicate that similarity does affect search times. In an elaboration of FIT, Treisman and Gormican (1988) proposed a Pooled Response Model to account for similarity effects.

The Pooled Response Model

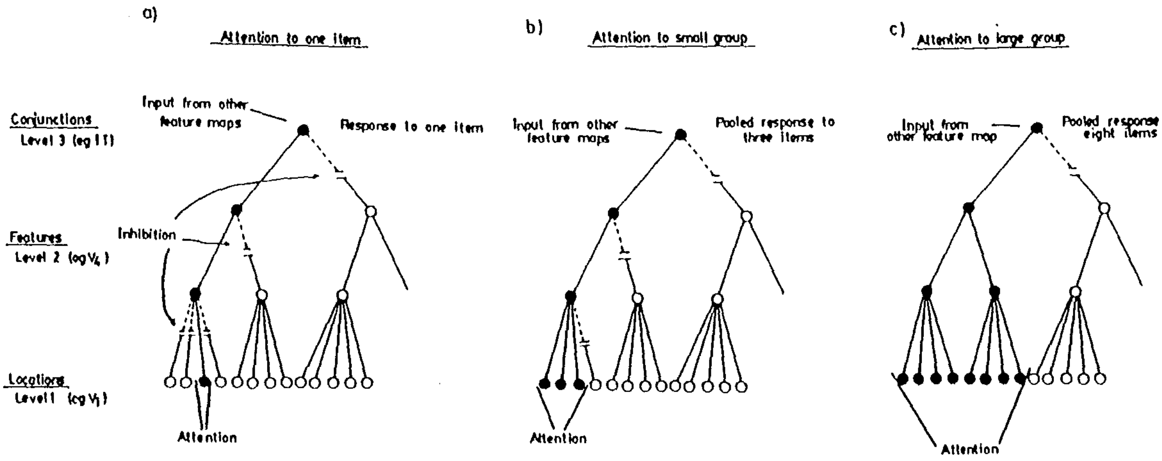
Treisman and Gormican added the Pooled Response model to FIT to account for two phenomena. The first is similarity effects on visual search. Treisman and Gormican (1988) found that serial search is needed to make very fine discriminations between stimuli that share the same feature but differ in the amount or size of that feature (e.g. two lines that differ in length). The model can also account for the fact that there are differences in search latencies dependent on the degree and type of similarity between targets and distractors. For example, Treisman and Gormican (1988) found that serial search was indicated by search latencies when the target was a shorter vertical line than a set of vertical line distractors. However, when the target was a longer vertical line than the distractors, parallel search was indicated. That is, they found a marked search latency asymmetry between targets that were either more or less of the same feature. They also found that when the line lengths were made less discriminable (i.e. the lines were more similar in length) the search functions for the longer line targets became much steeper, and very similar in slope to search functions for the shorter line targets.

The second phenomena the Pooled response model is proposed to account for is the discovery by Treisman and Souther (1985) that serial search was required for detecting the absence of a feature. Treisman and Gelade (1980) had proposed that search for feature targets should not give rise to the linear functions indicating serial search. However, a study by Treisman and Souther (1985) showed that although pop-out (or parallel search) occurred when the target was a circle with a radial line in distractors that were circles with no lines, the reverse condition did not produce pop-out. That is, a linear function indicated serial search was used by subjects to detect the absence of a feature when the target was a circle among distractors that were circles with radial lines.

In the Pooled Response model Treisman and Gormican (1988) propose that when someone is detecting visual stimuli, a pooled response from each relevant feature map is checked for the presence of activity anywhere in the map. The pooled response is hypothesized to be an average of the activity in the feature detectors. Once a filled location, as indicated by the location map, is attended to, activity from each feature map is restricted to only those features linked to the attended location (see Figure 4, page 21). With the addition of the Pooled Response Model, FIT can explain why serial search is required for the absence of a feature, as well as similarity effects on search latencies, in the following way: First, when attention is diffusely spread over a whole display (i.e. in the preattentive stage) only the presence of a unique feature will be detected and not its location (i.e. feature pop-out will occur). However, serial search and focal attention will be required if a target's defining feature is not unique, but also found in other items, as in searches for conjunction targets. The model also predicts that narrowly focused attention and serial search will be required for a target which is defined by an absence of a feature, as well as for features that differ only by degree along the same dimension. If search is for an item defined by the absence of a feature, as the circle without a line in Treisman and Souther's experiment, then there would be no activation of a detector, as only the presence of a feature would lead to activation. The model therefore predicts that serial search would be required to find the target. This would also be true for a target that was defined by having less of the same feature, such as the shorter line target in Treisman and Gormican's experiment, because a short line would lead to less detector activation than a longer line. Also, as differences in line length became less when the target was a longer line, activation would also become more similar to that for the distractor and more serially focused attention would be required in order to find the longer line. Treisman and Gormican suggest that Weber's Law with regard to just-noticeable-differences

(JNDs) would predict when attention and serial search would become necessary in order to detect the longer line target.

Figure 4: The model suggested by Treisman and Gormican "for pooled response and attentional control of feature selection" (copied from Treisman & Gormican, 1988, p46).



Problems posed for FIT by Visual Search Experiments.

Visual search evidence published subsequently to Treisman and Gelade's (1980) results has posed some problems for FIT. Some researchers have found that conjunction targets have exhibited pop-out (Humphreys & Bruce, 1989). One example is that Enns and Rensink (1990) found that three-dimensional cubes produced pop-out while two-dimensional items comprised of features contained in the cubes did not exhibit pop-out. A second example is that although Treisman and Gelade found that detecting an R in P and Q distractors required serial search (R has a conjunction of features found in the other two letters), Humphreys, Riddoch, and Quinlan (1985) found the effects of display size were small with an inverted T in homogeneous T distractors, even though the target contained both features of the distractors (a horizontal and a vertical line). However, both Enns and Rensink's, and Humphrey et al.'s findings could be explained by the "emergent features" proposed by FIT. The parts of a three-dimensional cube bear a certain relationship to each other which would not be apparent in the two dimensional items comprised of the same features. As Treisman and her colleagues point out, there could be "hard wired" detectors for such relationships. Also, although the inverted T in Humphrey et al.'s experiment contained the same features as the T distractors, these features bear a different relationship to each other in the target than they do in the distractors. There are, however, two other points to be made about Humphrey et al.'s experiment. There was only one type of distractor present in their experiment, while Treisman and Gelade's studies had two types of distractor. Also, when displays were regular shapes (i.e. the Ts formed a square or circle) responses on negative trials tended to be faster than on positive trials. These two points suggest the task of detecting the inverted T in T distractors could be more like a texture segregation task, as on negative trials there would be no initial break in the texture (Humphreys & Bruce, 1989). However, it still remains true that the only difference that would indicate a break in texture is how the features are related.

While the notion of emergent features might explain the apparently parallel search found in Humphrey et al.'s and Enns and Rensink's studies, it is more difficult to see how a result obtained by Nakayama and Silverman (1986) could be explained in the same way. These authors required subjects to detect colour-depth conjunction targets. Distractors were in two different colours at two stereoscopic disparities, so in near red and far blue distractors the target might be a near blue item. They found that search times did not differ much across different display sizes, indicating parallel search according to FIT. Humphreys and Bruce (1989) suggest that this may be because the display is initially coded into two depth planes and that once this occurs, the items on one depth plane would differ in only one salient feature (i.e. one near blue item in all near red items). The important point, however, is that even if Humphreys and Bruce's explanation is correct, a conjunction target was detected with apparently parallel search even when there are no obvious emergent features that might explain this.

Another problem for FIT is the fact that letters and words, although being made up of a number of constituent features, do not seem to require serial search to be integrated correctly (Humphreys and Bruce, 1989). This is apparent in Treisman and Gelade's (1980) experiments, in that even the letters used in the feature search condition actually comprise a combination of different form features. It is also apparent in a study by LaBerge (1983) when he found that subjects could attend to a word more easily than its constituent letters. However, FIT does allow that top-down processes might combine features in the absence of adequate focal attention. As letters and words are well learned one could expect top-down processing to effect their recognition. FIT does not, however, specify what these top-down processes are or how they might operate to combine features, and, as will be seen, there are other possible explanations.

Another problem posed for FIT by visual search experiments is that the reaction time functions that are seen to indicate parallel, preattentive search, are usually not entirely flat. Most visual search experiments produce a function that shows very slight increases in response time with increases in display size (e.g. Cohen & Ivry, 1991; Enns & Rensink, 1990; Nakayama & Silverman, 1986; Treisman & Gelade, 1980). Also, Prinzmetal, Presti and Posner (1986) showed that attention has an effect on the rates of feature errors as well as conjunction errors. In their experiments the stimuli were displays of four coloured letters. At the beginning of each trial subjects were first shown which coloured letter the target would be, and were also cued for target location. There were two conditions; a feature condition and a conjunction condition as in Treisman and Gelade's experiments. On some trials in each condition, the cue was invalid. The subjects' task was to report whether the target was present or not, and false alarm rates were compared for each condition and for valid and invalid cueing. Their results showed more false alarms occurred in the conjunction condition than in the feature condition. However, the cueing condition affected the error rates for both stimulus conditions and this was interpreted as indicating that attention (captured by the cue) affected feature detection as well as conjunction detection.

More recently, Cohen and Ivry (1991) have shown that the density of items in a visual scene affects visual search times. Although Treisman (1982) found no density effects, other researchers have found that the increasing density (or closeness) of items as display size (i.e. number of items) was increased, has affected search times (e.g. Pashler, 1987). Cohen and Ivry conducted two experiments in which subjects were required to indicate whether a target was present in each display or not. In one experiment targets were all conjunction targets (i.e. targets contained a conjunction of the features contained in the distractors). In the other experiment all targets were feature targets (i.e. the target differed from distractors by only one feature). In both experiments display size (i.e. number of items in the display) was varied over four sizes, and the spread of the display was varied to create two conditions; a "spread condition" in which items were separated by more than 1° of visual angle, and a "clump condition" in which the distance between items was 0.62° of visual angle. Cohen and Ivry's results indicated there was no significant difference in search latencies between the clump and spread conditions in the feature task experiment, but a significant difference in search times between these conditions was apparent in the conjunction target experiment. Search for the target was slower in the clump condition (when items were grouped closely together) than in the spread condition. Search was also slower for target absent trials than for target present trials. However, while the ratio for reaction time function slopes was 2:1 for target absent to target present trials in the clump condition (as Treisman and Gelade (1980) found in there visual search experiments), the ratio found in the spread condition varied from 1.5:1 to 1:1 over two experiments. As Cohen and Ivry (1991) point out, it is difficult for FIT to explain this result unless the focal attention, hypothesized as necessary for the correct identification of items comprised of a conjunction of features, could serially scan the visual field faster when objects are spread apart, than when they are grouped closely together. An attentional focus that moves serially from item to item across a visual field should take longer to scan a greater area than it would to scan a smaller area, as would be the case when items were clumped together.

Cohen and Ivry suggest that a two-stage location mechanism can explain their results. The first stage is proposed to be a coarse location mechanism that is fast and uses only course location information to bind features into objects, but can only operate when objects are not crowded together in a visual scene. The second is a slower, focal attention mechanism that is required when items are clumped together. However, it should be noted that Cohen and Ivry's hypothesis, like FIT, is unable to explain the slight increase in search times for feature targets as display size (i.e. item numbers) increase, apparent in their experiments as well as in previous studies.

A final problem for FIT in relation to visual search is posed by Townsend (1990). He points out that the visual search method cannot distinguish between a parallel system that is limited in capacity, and the serial system proposed by FIT to explain the linearly increasing reaction times apparent in the results for the conjunction condition targets. He points out that these linear functions indicate a limited capacity of some sort, but also that the limitation could "be due to seriality, limited capacity parallel, or even hybrid processing mechanisms" (p47). The possibility that items in a display in visual search experiments could be processed or attended to serially in small groups, rather than in a single item by item fashion, has been noted by Treisman as well as other authors (Humphreys & Bruce, 1989; Treisman, 1985, 1988). Townsend points out that a limited capacity parallel system may be a more parsimonious account of visual search results than the serial system suggested by FIT.

To summarize, there are four main problems posed for FIT by visual search experiments. First, targets comprised of a conjunction of features have exhibited pop-out. Although the possible emergent features and top-down processing proposed by FIT could explain some of the findings, these proposed processes do not easily explain other results, like those of Nakayama and Silverman (1986). Second, most visual search experiments show that search times for feature targets do increase slightly with increases in display size, and Prinzmetal et. al. (1986) demonstrated that attention does affect error rates in feature search as well as conjunction search. Third, the density of items in a visual display affects search times for conjunction search in a way that is difficult for FIT to explain. Cohen and Ivry (1991) have demonstrated that search is faster for conjunction targets when items are spaced more than 1° of visual angle apart than when they are grouped together so that they are separated by less than 1° of visual angle. Finally, Townsend (1990) has pointed out that, although Treisman and her colleagues have inferred serial processing from the linearly increasing reaction time functions found for conjunction search, it may be more parsimonious to assume that these functions reflect limited capacity parallel processing.

The Guided Search Model

Wolfe, Cave and Franzel (1989; see also Cave & Wolfe, 1990) have proposed an alternative to FIT, the Guided Search Model, in order to accommodate some of the problematic results posed by visual search experiments for FIT. Although the Guided Search Model does not offer a specific explanation of illusory conjunctions, I have included it in the thesis because it does specify how top-down processes might effect object detection. This factor could be an important consideration in explaining illusory conjunctions.

This model, like FIT, assumes a bottom-up analysis of the basic elements or features of the visual system. Also like FIT, the proposed first stage of analysis is the encoding of feature representations into feature maps. However, the Guided Search Model proposes that once features are registered in the feature maps, activation from those feature maps is then summed in the master map (called the Activation Map in this model) for each item location. This model also proposes that activation for features may be increased by top-down processes. So, in visual search experiments in which subjects have prior knowledge of what the target is, this prior knowledge will lead to increased activation of all features contained in the target. The summing of this activation in the Activation Map would lead to greater activation for the target location than for the locations of distractors, thereby guiding attention to likely target locations. For example, if the target were a red O in green O and red X distractors, then knowledge that the target is coloured red would lead to more activation in the colour map when red is detected than when green is detected. Also, there would be more activation in the shape feature map when O is detected than when X is detected. Therefore, when searching for a red O target, the sum of the activation (in the Activation Map) from the colour and shape feature maps at the location of the target would be greater than the summed activation at any of the other item locations. Figures 5a, 5b and 5c show diagrams for the processes described in the Guided Search Model.

Figures 5a, 5b and 5c: The Guided Search Model (copied from Cave and Wolfe, 1990).

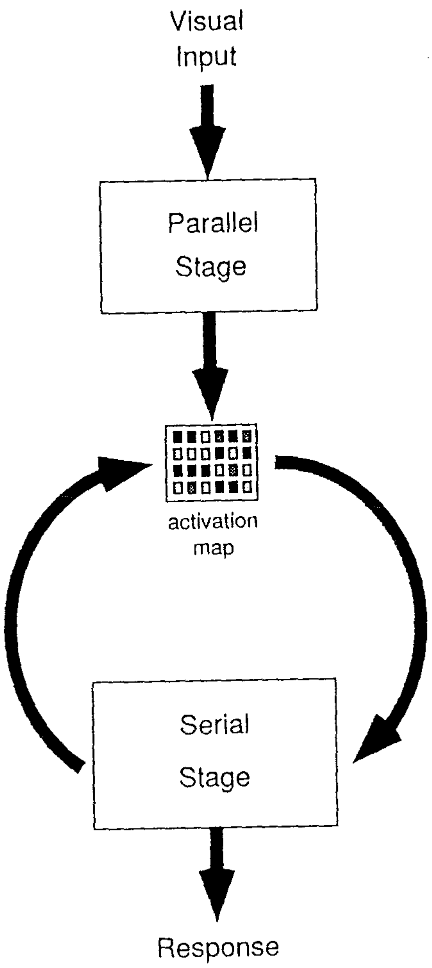


Figure 5a shows the operation of the model. "The parallel stage produces an activation map to guide the serial stage as it searches for the target. Each time the serial stage begins a new cycle, it begins to process the element with the highest activation in the Activation Map. When it determines that an element is not a target, it eliminates that element from further consideration" (p 231).

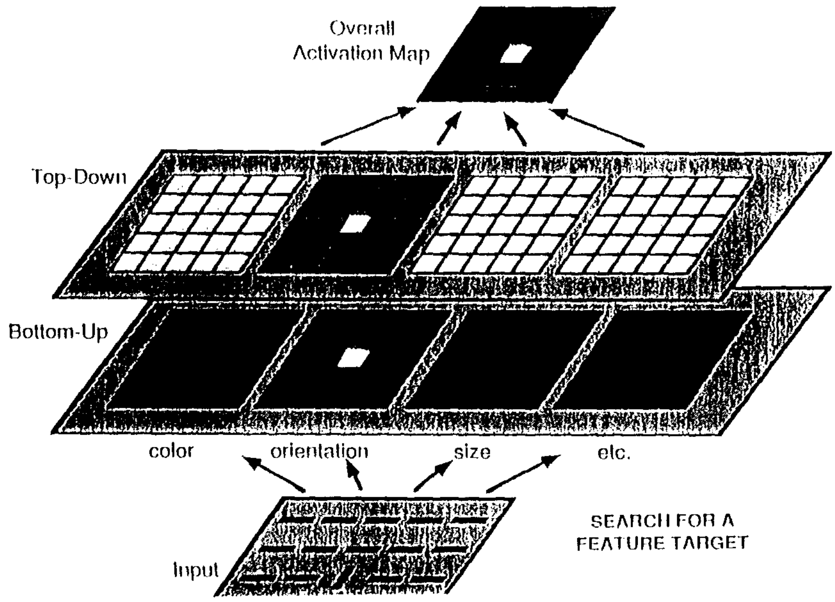


Figure 5b: "The operation of the parallel stage in a feature search with no noise present. Within each of the 3 layers (bottom-up, top-down, and overall), lighter shadings represent relatively higher activations, indicating a higher likelihood that a target is present at that location. In this feature search, the target has far more activation than the distractors in the orientation maps. When the maps are all summed to produce the Overall Activation Map, the target clearly has more activation than the distractors, and is always found quickly" (p234)

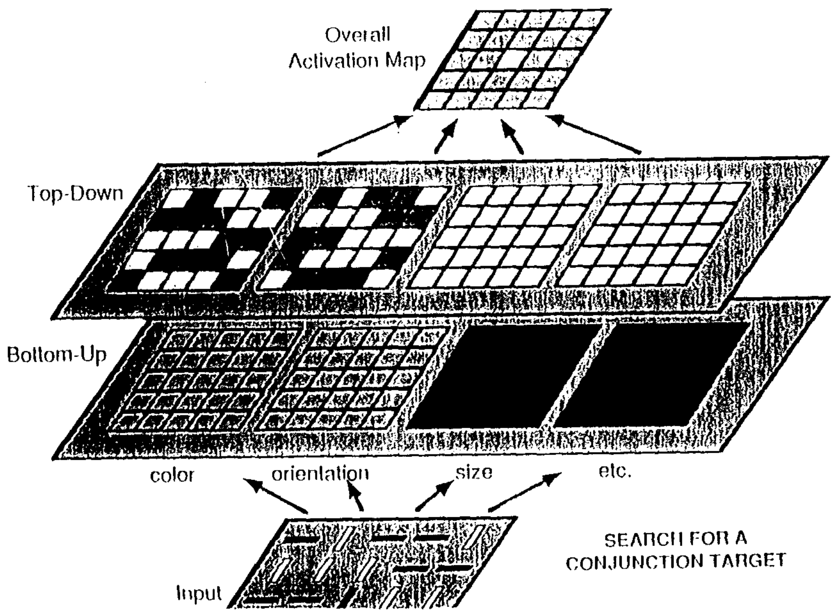


Figure 5c: "The operation of the parallel stage in a conjunction search with no noise present. As with the feature search, the target activation in the Overall Activation Map is still higher than the distractors, but the activation gap is much smaller. If noise is added, the target will sometimes be obscured" (p236-7).

While FIT suggests that the parallel and serial processes implied by visual search experiments are autonomous, the Guided Search Model assumes that information collected by the parallel system (i.e. the feature maps) would be available to the serial system, and thereby be able to aid serial search. This implies that completely serial exhaustive searches would not be required to find a conjunction target; a notion that is compatible with Townsend's (1990) view that linear increasing reaction time functions most likely reflect a limited capacity parallel system. The model predicts that factors other than just whether the target is defined by a conjunction of features found in the distractors would affect search latencies. For example, the salience of a particular feature might lead to a much greater difference in target activation over "noise" from the distractors, such that very fast search times could result even for conjunction targets.

Wolfe et al. (1989) also point out that, for their model, there is no qualitative difference between search for feature targets and search for conjunction targets. "In both cases the parallel processes provide the signal to guide attention to the target. In simple feature search the presence of a unique feature (e.g., red among green) generates a strong signal that quickly exceeds the background noise. As stimulus salience is reduced, the signal indicating the presence of a unique feature is hidden by the noise. In this case, the serial process would examine more of the distractors before the signal from the appropriate feature map exceeded the noise in that map and guided the serial process to the target location" (Wolfe et al., 1989, p428).

The Guided Search Model assumption that information from the feature maps can aid serial search, means that this model can accommodate visual search experiment results that FIT has difficulty explaining. First, because attention is guided by the summing of activation in the feature maps, the very fast search times producing almost flat functions for some conjunction targets can be explained. For example, the Guided Search Model is able to accommodate Nakayama and Silverman's (1986) results that indicated parallel search for the detection of colour-depth conjunction targets. It would also predict that some feature target detection tasks might produce slight increases in search times as display size increases, thereby explaining why this has occurred in many of the visual search experiments conducted (e.g. Treisman & Gelade, 1980; Treisman & Gormican, 1988).

This model could also accommodate Prinzmetal et al.'s (1986) finding that attention does affect the rates of feature errors as well as conjunction errors. Interestingly, the Guided Search Model could imply that attention is not necessary to, but rather facilitative of, combining feature representations into object representations: the level of activation, in the activation map for certain locations, indicates possible locations of the searched for object according to this model, therefore implying that attention is guided to the location only to ascertain whether in fact the object is situated there. That is, attention might increase accuracy by checking that the location does hold the target, and therefore facilitate correct identification by ascertaining that the activation level is indicative of the target and not just "noise".

Another result that the Guided Search Model can explain more easily than FIT is that some visual search triple conjunction targets are found faster than the double conjunctions used by Treisman and Gelade (1980). For example, Wolfe et al. (1989) used conjunctions of colour, form and size, rather than the colour and form conjunctions used in many visual search experiments. These triple conjunction experiments might, for example, require subjects to detect a big black vertical line in big white horizontal, small black horizontal, and small white vertical line distractors. While FIT predicts that search would be equally slow and serial for triple as well as double conjunctions, the Guided Search Model predicts faster search times if the target differs from the distractors by two features as it did in Wolfe et al.'s experiment, because this greater difference would lead to a greater difference in the activation from the target location than would be the case for double conjunctions. Wolfe et al.'s results were as their model predicted, with flatter functions being apparent than for double conjunction search. Where triple conjunction targets differ from distractors by only one feature, search times are very similar to those for double conjunctions, also as predicted by this model (e.g. Quinlan & Humphreys, 1987; Wolfe et al., 1989).

Although the Guided Search Model is able to explain some visual search results that FIT has been unable to accommodate, there are also some results that neither theory can explain. As Cohen and Ivry (1991) point out, the Guided Search Model, like FIT, cannot in its present form account for the density effects they found in their experiments, or for the fact that the 2:1 ratio for search times for negative versus positive trials does not hold for all conjunction searches. They also suggest that, with the addition of their two-stage location mechanism, the

Guided Search model could account for density effects. Because the Guided Search model does not offer an explanation of illusory conjunctions, for which density effects may have important implications, alternative explanations of density effects will be discussed in the section on illusory conjunctions and attention.

FEATURE LOCATION EXPERIMENTS

As well as conducting visual search experiments, Treisman and Gelade (1980) also conducted feature location experiments as part of the convergent evidence offered to support FIT. These experiments were to test the prediction that, while items comprised of a conjunction of features should require locating in order to be detected, features could be identified without necessarily being located. In these experiments subjects were briefly shown displays of two rows of six coloured letters, and the task was to report if a target was present, as well as its location. Distractors were pink Os and blue Xs, and the conjunction condition target was a blue O. The feature condition targets were a pink or blue H, or an orange O or X, so differed from the distractors by only one feature. Figures 6a and b show possible feature and conjunction target displays.

Figures 6a and b show possible feature and conjunction target displays for Treisman and Gelade's (1980) feature location experiment.

$$\mathbf{O}^{p}\ \mathbf{X}^{b}\ \mathbf{X}^{b}\ \mathbf{O}^{p}\ \mathbf{X}^{b}\ \mathbf{O}^{p}$$
$$\mathbf{X}^{b}\ \mathbf{O}^{p}\ \mathbf{O}^{\underline{b}}\ \mathbf{X}^{b}\ \mathbf{X}^{b}\ \mathbf{O}^{p}$$

Figure 6a: Conjunction target display where p = pink and b = blue.
The target's colour is underlined.

X^b O^p X^b O^p H^p X^b
O^p X^b O^p O^p X^b X^b

Figure 6b: Feature target display where the target H differs from all of the distractors by its form.

The results obtained in these feature location experiments addressed two conditional probabilities: the conditional probabilities of correctly identifying a target in correct compared to incorrect locations, and the conditional probabilities of a correct compared to incorrect location responses for incorrect target responses. Treisman and Gelade analyzed incorrect adjacent location responses separately to incorrect distant location responses because the probability for adjacent locations was greater than for distant locations: that is, a target could have only three adjacent locations (one on either side and one either above or below, depending on whether the target was in the top or bottom row). They also suggested the separate analysis of adjacent and distant locations could control for a bias to report locations adjacent to the target . As predicted, they found the probability of correctly locating compared to wrongly locating an incorrect target was not significantly above chance for feature or conjunction targets. FIT predicts that it should be possible to detect features without their locations, but not locations without target detection. Also as predicted, the probability of reporting a conjunction target correctly with an incorrect location response, was not above chance. FIT predicts that attention to the target location would be required in order to identify a conjunction target. In contrast, correctly identified targets without correct location did occur above chance for feature targets. This result was also as predicted. According to FIT, feature detection should be possible before items are located because feature information is encoded preattentively and separately to location information.

Problems posed for FIT by feature location experiments.

Johnson and Pashler (1990) point out a weakness in Treisman and Gelade's experimental design. Displays of rows of items could readily lead to reporting-errors occurring even though the target had been correctly located. They conducted similar experiments to Treisman and Gelade, but in order to control for reporting-errors they presented eight items in a hollow square, so that every item had a unique location (e.g. centre top, centre right, top right corner, etc). Figure 7 shows a diagram of Johnson and Pashler's display.

Figure 7: The geometric design of Johnson and Pashler's (1990) location experiment displays. Letter positions are shown with an X but actual displays used the same letters and colours as Treisman and Gelade's (1980) location experiment.

X X X
X   X
X X X

Johnston and Pashler further decreased the chance of reporting-errors by using the same geometry in the masking display, but used segments of colour rather than letters. They also controlled for guessing by including trials that did not contain a target, so subjects could report "no target" rather than being forced to guess. Their results indicated a close binding of the perception of identity and location for both feature and conjunction search. In fact, after correction for guessing, they found a weak trend for accurate perception of location without correct identification of the feature target. Johnston and Pashler's results suggest that in order to identify even simple features, they must be located, or at least that feature and location information become available in parallel. This is counter to the FIT predictions that location is necessary only for identifying conjunction targets and that features might be identified without location information. The Guided Search Model can, however, accommodate Johnston and Pashler's results because it predicts no qualitative differences between feature and conjunction target search. "In both cases the parallel processes provide the signal to guide attention to the target" (Wolfe et al., 1989, p428).

TEXTURE SEGREGATION

To test the prediction that features should mediate "preattentive" and easy texture segregation while a conjunction of features should not, Treisman and Gelade (1980) measured the amount of time it took subjects to sort cards according to where they detected a division, or boundary, in a matrix of five rows by five columns of items. On some cards the division was specified by a difference of one feature (shape or colour) between two groups of items (e.g. a group of red Os and Vs with a group of blue Os and Vs). On other cards the division resulted from a difference in a conjunction of features (e.g. a group of red Os and blue Vs with a group of blue Os and red Vs). The results were consistent with Treisman and Gelade's predictions. Sorting times were significantly slower for the conjunction condition (a mean of 24.4 seconds) than for the feature conditions (14.5 and 16.2 seconds for the colour-feature and shape-feature conditions respectively). In another experiment these authors obtained similar results using local components of shape rather than values on different dimensions, like shape with colour. For example, EO with FQ in the feature condition, as these letters differ by only one feature, and FQ with EX in the conjunction condition, where no simple features distinguish the letters from each other. Treisman and Gelade concluded that the critical variable determining automatic texture segregation was whether areas differed in a single feature or a conjunction of features, and further, that automatic texture segregation requires the preattentive detection of homogeneities.

Texton Theory.

Treisman and Gelade's view of texture segregation has much in common with Bela Julesz' Texton Theory. Julesz defines textons as local conspicuous features, combinations of which all textures can be reduced to (Julesz, 1975, 1981,1984; Julesz & Bergen, 1983). Initially he postulated only three types of textons: elongated blobs such as rectangles and ellipses, terminators of line segments, and crossings of line segments. However, he has more recently included binocular disparity, movement parallax (or velocity) and flicker as possible textons (Julesz, 1984). Like Treisman's FIT, Julesz' Texton theory postulates preattentive and attentive stages of vision, and also that the positional relationships of features are ignored in the preattentive stage. He suggests that "preattentive vision directs attentive vision to the locations where differences in textons or in density of textons occur" (Julesz, 1984, p42).

Texton theory and FIT also diverge on some important points. The relative salience of features has been found to affect texture segregation performance (e.g. Callaghan, 1989; Enns, 1988; McIlhagga, Hine, Cole, & Snyder, 1990; Pashler, 1988). While Texton Theory takes no account of the relative salience of features, the Pooled Response Model, added to FIT by Treisman and Gormican (1988), can account for this phenomenon. For example, Enns (1986) found that relative line lengths affected discrimination accuracy more than differences in texton type or number, and that closure could facilitate the segregation of areas; results that Texton Theory has difficulty accommodating but that can be accounted for by FIT and the Pooled Response Model. As Enns (1986) points out, his results suggest that features, (or textons, or the basic elements of visual cognition) are graded, rather than entirely categorical as Julesz suggests; a point that FIT does acknowledge.

Problems posed for FIT by texture segregation experiments.

There are, however, problems associated with texture segregation experiments for FIT as well as Texton Theory. Visual search results have indicated attentive serial search for some items that texture segregation experiments have indicated preattentive parallel processing for (e.g Treisman & Souther, 1985). Enns (1986) suggests that differences in experimental procedure, as well as the relative salience and graded nature of feature representations, could explain the different findings. However, which particular aspect or aspects of the procedures are responsible for the different findings is an important issue. In visual search experiments items are distributed randomly with one target item placed among those items, so elements of the display are *separate whole objects*. In texture segregation experiments homogeneous groups of items are assigned to adjacent areas, and subjects are required to find the boundary between the groups, so elements of the display are a *part of the surface of an object or area* and further, the subjects' task focuses on their use of processes that might be used to separate objects or areas of space from one another (i.e. the boundaries between areas of space). It is possible that these global aspects of the visual presentation are in part responsible for the different results, because they lead to different types of processing. Given that the tasks required of the subjects are different in the two procedures, it seems likely that the elements of the different displays are processed differently in each of them. While FIT can account for graded elements giving rise to texture segregation, like relative line length, and Texton Theory cannot, neither theory can easily account for the possible effects of the more global aspects of the visual displays in these experimental procedures. Cohen and Ivry's (1989, 1991) two-stage binding and location mechanism might account for these effects, as it can account for density effects. So also could Allport's (1989) suggestion that visual attention could be organized by constraints such as the Gestalt principles like proximity, number and closure. I will return to these possibilities in the section on illusory conjunctions and attention.

ILLUSORY CONJUNCTIONS

Treisman and Schmidt (1982) added the final piece of convergent evidence to support FIT, when they tested the prediction that features could be incorrectly combined to produce illusory conjunctions in the absence of focal attention. Illusory conjunctions are errors made in visual cognition that combine features from two separate items in a visual display, so they are perceived as a coherent object that is not actually present. For example, a report of a green X when the visual stimuli were actually a pink X and a green T. Treisman and Schmidt presented subjects with displays of three coloured letters between two black digits. The primary task consisted of reporting the two digits as a single number (e.g. 6 and 4 as 64). A secondary task produced the results of interest, and consisted of reporting anything else confidently observed. Figure 8 shows a possible display used in this experiment.

Figure 8: A possible display for Treisman and Schmidt's (1982) illusory conjunction experiments, showing the digits placed on either side of the 3 coloured letters. The colours of the letters are shown beside the letters.

6 T pink **S** blue **N** green **4**

Three types of error could occur for the coloured letters: feature errors which combined a correct feature with a feature not present in the display; illusory conjunctions which combined two features from different items in the display; and complete errors, in which neither of the features reported were present in the display. The coloured letter displays were derived from five letters and five colours. Because there were three letters in each display, illusory conjunctions and feature errors would have an equal chance of occurring (i.e. a ratio of 1:1 illusory conjunctions to feature errors would be the chance level). For example, if a letter was reported correctly and its colour incorrectly, then there would be two colours present (in the other two letters) in the display that would result in an illusory conjunction if reported, and two colours not present in the display which would lead to a feature error if reported. Treisman and Schmidt found that illusory conjunction errors significantly exceeded feature errors, the latter resulting when subjects reported one correct feature with a feature not present in the display. The excess of conjunction errors over feature errors was large, far exceeding chance (0.39 compared to 0.15) indicating that it was unlikely that all the illusory conjunctions reported were simply feature errors.

Treisman and Schmidt, utilizing recognition and simultaneous matching tasks, showed it was unlikely that illusory conjunctions resulted from a response bias (i.e. a bias to *report* conjoined features rather than report features separately), verbal label switching, or memory failures. In an experiment to address these issues, displays were the same as in the first experiment, but no display contained feature repetitions. For each display there were also three probe displays: a feature probe that combined one feature from the display with one not present; a conjunction probe comprised of two features present in different items in the display; and an identical probe that was exactly the same as one item in the display. Subjects were shown a probe immediately before each display and required to report the two digits, and then whether or not any coloured letter matched the initial probe. Using probes meant that subjects did not have to report what they had seen, merely respond yes or no. This could control for response biases, verbal label switching and memory failure. As predicted, significantly more conjunction errors than feature errors were made, but the difference was less than was apparent in the first experiment. Treisman and Schmidt suggest that this weaker result was most likely to reflect a contribution from memory failure in the recall task used in the first experiment.

In a third experiment, Treisman and Schmidt further reduced the possibility that verbal coding may have led to illusory conjunction reports. Displays in this experiment contained five coloured letters, four of which formed the corner points of a square, and the fifth one in the centre. As in the first two experiments, each of two black digits was displayed on either side of the coloured letters. Some displays contained an identical pair of coloured letters while others did not, and the subjects' secondary task was reporting whether a display contained at least one identical pair of items. Although, once again, the difference was smaller than in the first experiment, significantly more conjunction errors than feature errors occurred in this matching task.

As well as producing evidence that illusory conjunctions of features might be formed in the absence of focal attention, Treisman and Schmidt also further elaborated FIT by suggesting that prior experience and learning, via top-down processes, could explain why we do not regularly experience illusory conjunctions in everyday visual experience. They use the example of people not experiencing perceptions of a blue sun in a yellow sky even though their surroundings are often not attended to (perhaps not the best example as viewing the sky would not require the resolution and finely focused attention needed to detect a small coloured letter). This hypothesis suggests that top-down (e.g. memory) and bottom-up processes (e.g. automatic feature mapping) interact in the second of the three stages of visual detection proposed by FIT (Treisman, 1985). That is, once feature representations are activated in the feature maps they might be conjoined to form correct object representations by focal attention falling on the location of the objects, or by top-down processes in the absence of such attention. Although Treisman and her colleagues do not specify what the mechanisms of this top-down processing are, they could explain a problematic result for FIT obtained by Virzi and Egeth (1984), who found that illusory conjunctions can arise from more than just early visual processes. These authors used the same procedure as Treisman and Schmidt, but presented subjects with coloured words rather than single coloured letters between the two black digits. Some of the words actually named colours, so BIG, BLUE and WIDE might be presented coloured red, green and yellow respectively. Their results showed a greater than chance reporting of illusory conjunctions of words actually present with colours named but not present in the display (e.g. BIG reported as blue). As Humphreys and Bruce (1989) point out, this result contradicts the notion that only very primitive properties or features are integrated incorrectly in the absence of focal attention. However, it does not seem to contradict the notion that top-down processes may combine features when objects are not attended to. It is also possible that Virzi and Egeth's results might have eventuated from a response bias created by word naming a colour.

Other subsequent research to Treisman and Schmidt's has added further support to FIT's predictions regarding illusory conjunctions. This research has shown that illusory conjunctions can also be formed from two form elements as well as from the form and colour elements used by Treisman and Schmidt. For example, Treisman and Paterson (1984) found that an S and a vertical line could be erroneously conjoined by subjects to form illusory dollar signs. Also, Prinzmetal (1981) found that subjects reported illusory plus signs when vertical and horizontal lines were presented separately, each inside a circle. However, subsequent illusory conjunction research to Treisman and Schmidt's has also posed some problems for FIT. These problems, along with the relevant research and various explanations of illusory conjunctions, will be discussed in the next section.

4

ILLUSORY CONJUNCTIONS AND ATTENTION

To recap, Treisman's FIT explains the phenomenon of illusory conjunctions in the following way: The theory proposes that attention is necessary for correctly conjoining preattentively encoded, and spatially free-floating features, into whole coherent objects. If objects fall outside of the spread of attention, or inside of the spread but with insufficient time to attend to items fully, then incorrect conjoinings of features in the form of illusory conjunctions are predicted. Treisman and Schmidt, in conducting their experiments, made other important assumptions. They assumed a "zoom lens" model of attention and also that the primary digit task formed a boundary which the attentional lens was narrowed down to. Also, the theory suggests that illusory conjunctions could be formed both outside of the attentional boundaries (where features would be "free floating" in the absence of attention), as well as inside those boundaries when viewing time is very brief. Recall that the theory proposes that items falling outside of the attentional boundaries are processed by the feature maps only (i.e. the parallel system) but information from items falling inside the attentional boundaries is made available to the serial system where the conjoining of features into objects occurs. Because of this different processing of items falling inside and outside of the attentional boundaries, FIT also predicts that illusory conjunctions will not be formed from the features contained in items from both sides of the boundary: that is, an illusory conjunction could not be formed by conjoining a feature from an item outside of the focus of attention with one from an item that falls under the attentional "spotlight."

As well as demonstrating that illusory conjunctions are experienced above chance levels when viewing times are very brief, Treisman and Schmidt also made some other important observations. First, that the features, in the form of colours and letters, most often switched positions so that they did not appear to be experienced as duplications of themselves. In fact both feature switching and feature duplication have been observed by other experimenters (e.g. Prinzmetal, Treiman, & Rho, 1986; Prinzmetal and Keysar, 1989). Secondly, they observed that illusory conjunctions were no more likely to be formed from features contained in items that were close together than from those contained in items that were far apart. This observation further supported FIT which proposes that features are "free-floating" or spatially unlocated until they fall under the attentional "spotlight." However, evidence from other researchers has proved contrary to Treisman and Schmidt's observations, which poses a problem for FIT (e.g., Cohen & Ivry, 1989; Prinzmetal, Treiman, & Rho, 1986).

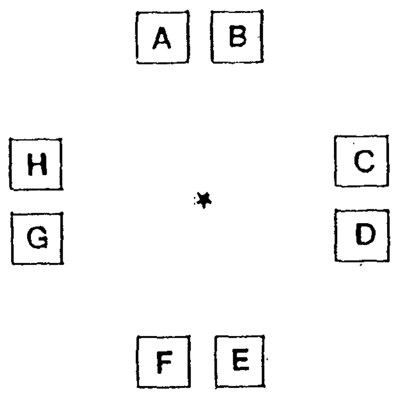
DISTANCE EFFECTS ON ILLUSORY CONJUNCTIONS

Evidence of the distance between items having an affect on the formation of illusory conjunctions would have important implications for FIT because the theory proposes that features are initially registered without location information. That is, features are encoded preattentively and are unlocated or "free floating." If the amount of separation between items effects the rate of illusory conjunctions, the implication is that there is at least some location information registered with, or even before, the features.

Cohen and Ivry (1989) found that the amount of separation between objects did have an effect on illusory conjunction rates when they conducted research that specifically addressed the effects of distance on the formation of illusory conjunctions. They presented subjects with an asterisk as a central fixation point followed by a white digit at the fixation point, which subjects reported as their primary task, and two coloured letters positioned in two of eight possible locations. The eight letter positions consisted of four location pairs placed so that one location pair was above, one below, one to the left and one to the right of the fixation point. There were therefore two top adjacent, two bottom adjacent, two left adjacent, and two right adjacent locations. By placing letters in positions around the primary task digit, Cohen and Ivry reasoned that, according to FIT, if the digit task captured a subject's focal attention then the letters would fall outside the attentional boundaries rather than inside as in Treisman and Schmidt's (1982) experiments.

In this experiment, the two coloured letters were presented either in adjacent positions (e.g. both in the top position) or in distant positions (e.g. one in the top position and one in the left position). This effectively created two conditions: an adjacent condition and a far condition. The visual angle between letters in the adjacent condition was less than 1° while the visual angle between the letters in the far condition was more than 2.5°. The subject's secondary task produced the results of interest, and consisted of reporting which of two possible target letters was present, an F or an X, and its colour. The other letter presented, the distractor, was always an O. Figure 9 shows how displays in Cohen and Ivry's experiment were arranged.

Figure 9 shows the possible location of the stimuli used in Cohen and Ivry's experiment. The star indicates the location of the digit (copied from Cohen & Ivry, 1989, p652).

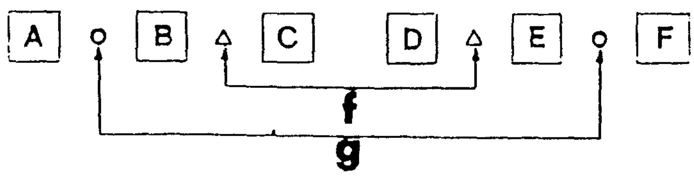


In this procedure, feature errors are indicated by subjects reporting the letter incorrectly with the correct colour, or the correct letter with a colour not present in the display. Illusory conjunctions were indicated by a correct report of the letter with the colour of the distractor O and a ratio of conjunction errors to feature errors greater than 2:1 indicated that illusory conjunctions were occurring above chance level. This ratio indicated chance levels of illusory conjunctions because target and distractor letters were varied over four colours. On any trial one of these four colours would be contained in the target and if reported would be correct. This would leave three possible error colours, one of which would be the colour of the distractor. If the distractor colour was reported in error this would constitute an illusory conjunction (i.e. the colour feature from one item conjoined with the form of the other item present in the display). The other two colours would not be present in the display so if either were reported they would constitute colour feature errors. This means that a chance rate of illusory conjunctions would be one for every two colour feature errors. Cohen and Ivry found that illusory conjunctions were reported significantly more often in the adjacent condition than in the far condition, and that their occurrence exceeded chance levels in the adjacent condition (mean proportional rates of 0.171 colour feature errors to 0.135 illusory conjunctions), but were significantly below chance level in the far condition (0.160 colour feature errors to 0.061 illusory conjunctions); a clear indication of distance effects on the occurrence of illusory conjunctions.

Cohen and Ivry conducted other experiments using the same method as the first one (and for which they did not give the mean error rates), but varied the visual angles between the adjacent letters. They found that when 1.5° of visual angle separated the letters, illusory conjunctions did not differ from chance levels. Also, if the visual angle between letters was increased to 2.17° or more, illusory conjunctions occurred significantly below chance levels.

In another set of experiments Cohen and Ivry adapted Treisman and Schmidt's (1982) procedure by using the same tasks as in the previous experiment, but presenting the two coloured letters (a target F or X and a distractor O) in two of six positions that were located in a row. The primary digit task consisted of reporting two digits positioned to the left and right of the display, and like Treisman and Schmidt they assumed the digit task formed the boundaries of attention. Based on this assumption, they varied the positions of the digits in order to manipulate the attentional boundary. In a "small spotlight" condition the digits were placed one on either side of the two central letter positions, while in a "large spotlight" condition they appeared on either side of the central four letter positions. Figure 10 shows how the stimuli were organized in these experiments.

Figure 10 shows the possible stimuli positions in the displays used in Cohen and Ivry's second set of (1989) experiments. Here f indicates the position of the digits for the "small spotlight" condition and g the position of the digits for the "large spotlight" condition (copied from Cohen & Ivry, 1989, p657).



Results for the small spotlight condition indicated conjunction errors occurred above chance level for only the central letter pair (i.e., when both letters appeared between the two digits). No other letter pairs showed significant levels of conjunction errors (i.e., when one letter was between and one outside of the digits, or both were outside and on either side of the digits). In the large spotlight condition significant levels of illusory conjunctions occurred for the central letter pair but not when the letters were placed further apart in the locations on either side of the central two positions, but inside of the digits (i.e both letters were inside the digits but separated by the two central letter locations). The level of illusory conjunctions when the letters appeared on either side of the digits was not significantly above chance, but there was a significant trend to report illusory conjunctions when one letter appeared inside (and adjacent to) and one outside of (and adjacent to) the digits: that is, when the letters appeared close to and on either side of one digit.

Together with the results for their first experiments, these findings raise some important points. First, the location of the letters, both in the distance between them and in whether they both fell inside the division created by the digits or not, seems to have considerable effect on whether illusory conjunctions occur at a significant level or not. As a result of adjusting the distance between the letters in the first set of experiments, Cohen and Ivry concluded that significant levels of illusory conjunctions were formed outside of attention only if the letters were separated by less than 1° of visual angle, and that a distance between them of more than 2.17° of visual angle resulted in fewer illusory conjunctions than would be expected by chance. In the last experiment described, illusory conjunctions were formed from letters appearing between the digits and adjacent to each other when they were separated by more than 1° of visual angle (i.e. agout 1.4°). So it also appears that illusory conjunctions are formed from letters which are further apart when thwy are centrally located. than would be the case when they are located peripherally. As Cohen and Ivry point out, FIT is unable to account for these distance effects, and in fact predicts that the amount of distance between items will not effect whether illusory conjunctions will be formed from their constituent features or not. Neither can FIT account for the trend to form illusory conjunctions when the letters in the latter experiment were presented on either side of, and adjacent to, one of the digits. Cohen and Ivry suggest that a two-stage location mechanism in conjunction with FIT could explain the distance effects.

The Two-Stage Location Mechanism.

In describing their two-stage location mechanism Cohen and Ivry, like Treisman and her colleagues, assume that attention is necessary for a "binding mechanism" that conjoins initially encoded features; that is, focal attention is used to mark an area, the boundaries of which are specified by the boundaries of the focus of attention, and then the binding mechanism "glues" together the features contained within that area. However, they propose that when features are initially encoded that some course location information is also encoded, rather than the features being entirely "free floating" as FIT suggests. Cohen and Ivry point out that the evidence from their experiments suggests that while the course location information would not be enough to prevent the mislocation of features to nearby locations, it would prevent their migration to more distant locations.

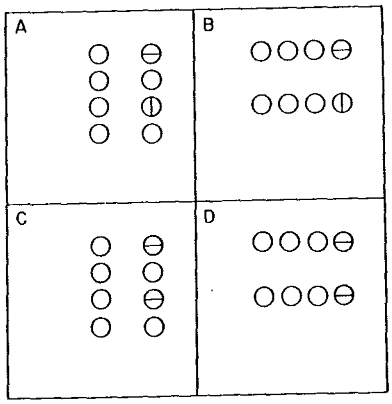
Assuming that attention constitutes the glue that binds feature representations into object representations, Cohen and Ivry point out an unresolved question regarding this proposed attentional binding mechanism: it is not known what factors might determine the boundaries of attention. Recall that FIT assumes attention has clear boundaries and, by implication, that the attentional boundaries constrain feature integration so that the features of objects falling under the attentional focus are combined while those features outside of attention remain free floating unless combined by top-down processes. Recall also that Allport (1989) suggested the Gestalt principles could constitute the constraints on attention. Cohen and Ivry propose a way in which this might be done by suggesting "that some of the Gestalt grouping principles may serve as cues for determining the boundaries" (Cohen & Ivry, 1989, p661) of the attentional focus. They also observed, following Marr (1982) who has noted the importance of edges in segregating visual space into objects, the possibility that the edges of objects might also determine the boundaries of the attentional focus.

An issue discussed in the section on visual attention poses a problem for Cohen and Ivry's suggestion that grouping principles or edges of objects might serve as the boundaries to attention: that is, research has indicated that the attentional focus may be more like a gradient and have no clear boundaries (Downing, 1988; LaBerge & Brown, 1989). Another problem for their hypothesis is, of course, that FIT's assumption of attention being the mechanism necessary to binding features into objects can be disputed. For example Tsal (1989) proposes the convergent evidence supporting FIT does not support the hypothesis that attention is necessary for combining features into whole objects. He also suggests that illusory conjunctions may not represent a function of non-attention. Tsal points out that, although the primary task in the illusory conjunction experiments has hypothetically led to a condition of non-attention for the coloured letter items in the displays, a very large proportion of these letters were still reported correctly. Like some visual search experiments this suggests that some features may be correctly combined in the absence of attention and without any obvious top-down processes. Tsal therefore proposes that the evidence as it stands is not consistent with the notion of attention being the sole factor responsible for feature integration, but rather, there may be other factors contributing to their correct integration. He notes a possibility that attention may facilitate the correct integration of features by improving their localization. This notion, that attention may be merely facilitative rather than necessary for feature integration, is also suggested by Prinzmetal and Keysar (1989) in an alternative explanation to FIT of illusory conjunctions. This alternative, along with the ensuing research, will be discussed in the following section.

GROUPING EFFECTS ON ILLUSORY CONJUNCTIONS

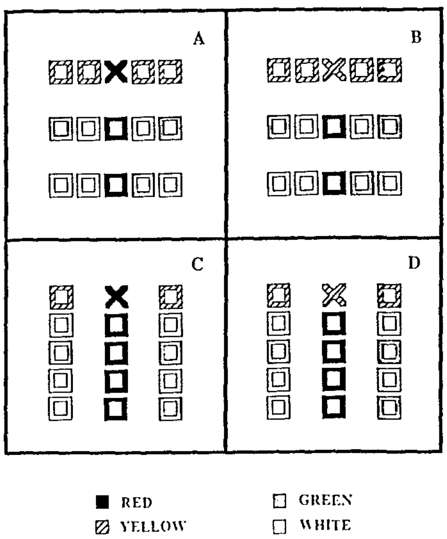
Prinzmetal (1981) found that illusory plus signs were reported when subjects were briefly presented with horizontal and vertical lines, with each line situated inside a small circle. The stimuli were, however, presented in two parallel lines of four circles, and Prinzmetal found that the illusory plus signs were only formed from two lines appearing within the same group of circles, not when each of the lines were presented in a different group. Figure 11 shows examples of displays used by Prinzmetal (1981).

Figure 11 shows 4 examples of displays used by Prinzmetal in his experiment (copied from Prinzmetal, 1981, p332).



Based on Prinzmetal's (1981) results, Prinzmetal and Keysar (1989) proposed a functional explanation of illusory conjunctions that did not include attention as a necessary factor in feature integration. Their explanation is based on two assumptions. First that there is poor spatial resolution for some aspects of visual information and second, that spatial information is constrained by perceptual organization. They propose that these constraints result from cognitive processes that cause perceptual grouping and divide visual space into areas and objects. These processes would reduce spatial uncertainty (perhaps providing the course location information suggested by Cohen and Ivry to explain their results) and would constrain feature integration so it did not occur across the boundaries of objects or areas of space. Prinzmetal and Keysar point out that their explanation is functionally adaptive in terms of cognitive economy, and suggest that illusory conjunctions are the result of this functionally adaptive economy. They suggest that when visual perception opportunities are limited or very brief, this would lead to poor or degraded information for features and items contained within the areas defined by the constraints. They further suggest that a spreading effect from the information that is registered would fill in the missing data, thereby leading to illusory conjunctions. While Prinzmetal and Keysar's theory does not hold attention as necessary for the combining of feature representations into object representations, it does, like FIT, assume that visual attention is spatially based (i.e. the early attentional selection hypothesis), and predicts that attention can affect feature integration by improving both location and feature identity information.

Figure 12 shows 4 examples of displays used in Prinzmetal and Keysar's experiment (copied from Prinzmetal & Keysar, 1989, p171).



Prinzmetal and Keysar used a computer simulation of the mechanisms suggested by their theory to demonstrate that these mechanisms are sufficient to produce illusory conjunctions. They also conducted a series of experiments to test the predictions made by their theory. In one experiment subjects were briefly presented with clearly defined groupings of items in either three rows or three columns of five coloured letters. Subjects were required to report the colour of a target X. Figure 12 shows examples of displays used in their experiments.

Prinzmetal and Keysar's results demonstrated that significant levels of illusory errors that attributed the colour of an adjacent letter in the same row or column as the X occurred, but no significant level of errors that attributed the colour of an item in another row or column occurred. There was no primary task included in this experiment that might have hypothetically restricted the focus of attention to only one group of items, and subjects had to find the randomly placed target anywhere in the display.

In another experiment these authors adapted Treisman and Schmidt's (1982) procedure, using the primary digit task with an evenly spaced matrix of four by four items. On some trials the digits were placed vertically, above and below the matrix, while on other trials they were placed horizontally on either side of the matrix. The results showed that significant levels of illusory conjunctions occurred only along a row that contained the target X when the digits were placed horizontally, and only within the column containing the target when the digits were placed vertically. Prinzmetal and Keysar therefore demonstrated that the subjective organization or grouping of visual information could affect where illusory conjunctions occurred. In effect, they demonstrated how Treisman and Schmidt's (1982) as well as Cohen and Ivry's (1989) results might be explained by their theory.

For example, the manipulation of the digits in Cohen and Ivry's experiments could have led to a different perceptual grouping rather than to a shift in the boundaries of attention.

Other research also lends considerable support to Prinzmetal and Keysar's hypothesis. Ramachandran's (1990) motion capture experiments discussed in the section on visual attention demonstrated how phenomena that created a subjective experience of a boundary led to all the items within that boundary being grouped and treated by the visual system in the same way. That is, apparently moving gratings and illusory squares captured the dots contained within them into their apparent motion. More recently Prinzmetal, Hoffman and Vest (1991) found that syllable-like units, defined by orthographic patterns and morphemes, affected the rate of illusory colour-letter conjunctions reported by subjects when they were required to report the colour of a target letter contained within a word. Subjects were more likely to report the colour of another letter contained within the same unit than the colour of one contained in another unit of the same word.

Prinzmetal and Keysar's experimental results raises particular difficulties for FIT. Their subjects were not given an attentional cue and the target, when one was present, could appear randomly anywhere in the displays. As Prinzmetal and Keysar point out, FIT could only somewhat implausibly, explain the results by supposing that the attentional focus was, in the brief presentation time, narrowed down to either the row or column containing the target. This might explain why illusory conjunctions were not formed from features from two different groups of items by assuming the features did not then cross the attentional boundaries. However, it leaves open the question as to why attention would select the whole line of items rather than only the immediate target area.

CONCLUSIONS

Prinzmetal and Keysar's results appear to indicate that attention and its hypothetical boundaries are not the key issue in the formation of illusory conjunctions; that rather illusory conjunctions result from adaptive constraints on visual processing. Their theory and results also suggest a hierarchy of events in visual processing, and more particularly in feature integration; a hierarchy similar to the one suggested by Cohen and Ivry (1989,1991) in their proposed two-stage location mechanism. It appears that a set of constraints might first determine the separation of visual space into areas and then the visual information contained within those areas is processed in "chunks" or units. Prinzmetal and Keysar's (1989) results suggest that both subjective and objective groupings could constitute some of those constraints. This notion is supported, and other likely constraints suggested, by some of the problematic results from the other experimental paradigms used in obtaining the convergent evidence to support FIT.

First, Nakayama and Silverman's (1986) results could be explained by the notion that the visual scene is first divided into areas or groups by certain constraints. Their results indicated parallel search when subjects were required to detect colour-depth conjunction targets. It is possible that visual scenes are initially divided into different depth planes (Humphreys & Bruce, 1989; Treisman, 1988); that is, that binocular disparity cues constitute some of the constraints on visual processing. There is neurological evidence to support this notion. For example, cases of unilateral neglect following injury confined to only distant objects (Bisiach, Perani, Vallar & Berti, 1986; in Allport, 1989), and experimental demonstrations of unilateral neglect confined to near objects in monkeys (Rizzolatti, Gentilucci & Matelli, 1985; in Allport, 1989). If the visual scene is initially coded according to different depth planes then the subjects in Nakayama and Silverman's experiment would have been confronted with only a simple feature detection task (i.e. finding a near blue target in near red distractors rather than in the whole set of near blue and far red distractors).

The notion of initial constraints, particularly the Gestalt grouping principles, can also explain the different results obtained for the same stimuli in texture segregation and visual search experiments. It may be true that the factor which differentiated one texture area from another in Treisman and Gelade's (1980) and Julesz' (1974, 1981) texture segregation experiments was that items from each area differed by only one feature. However, results from visual search experiments indicating serial search for some items, that texture segregation experiments indicate preattentive parallel search for, suggest other factors may also be involved in texture segregation. Cohen and Ivry's (1991) density effects on visual search indicate that other factors are likely to be the density and proximity of items in the visual field.

Another visual search result previously mentioned as a problem for FIT, was a finding of Humphrey et al. (1985). They found that the effects of display size were small for an inverted T target in a set of T distractors. In these experiments the items were presented in close proximity to each other, so density may have contributed to faster parallel processing. Further, they found that when displays were in regular shapes (i.e. the Ts were presented in a group that formed a square or circle) that responses on target absent trials were faster than when they had not been grouped in such a fashion, and also faster than for target present trials. Cohen and Ivry (1989) suggested that object boundaries as well as grouping might provide initial or first stage location information. Perhaps the shape boundary as well as the close proximity of the items in Humphrey et al.'s experiments facilitated faster processing of the visual information.

Prinzmetal and Keysar's (1989) theory and their experimental results are also compatible with Allport's (1989) suggestion that visual phenomena, such as the Gestalt grouping principles, could constitute the constraints on visual attention. Logically, if certain visual phenomena constrain attention they would be encoded before attention was applied to any particular item in a visual scene; that is, some visual processing would be required in order to make available the information by which attention is constrained or ordered. Allport's notion of visual attention as an adaptive selection-for-action mechanism is also conducive to the idea that it would be adaptive to initially encode spatial information that divided visual space into the possible areas and objects with which an organism might interact.

To summarize, the evidence from illusory conjunction experiments, as well as some visual search experimental results that have posed problems for FIT, support the notion that information dividing visual space into areas or units is initially encoded and then constrains *both* visual attention and feature integration. The evidence also suggests that this initially encoded information not only limits or constrains feature integration, but that it may, under some conditions, also facilitate the processing of visual elements contained within the areas or units of space, by enabling more cognitively economical use of a limited capacity parallel system. It further suggests illusory conjunctions are the consequence of a spreading effect, from the degraded visual information registered as a result of limited exposure time. Finally, the evidence indicates that visual attention is not necessary to feature integration, but rather, facilitates the correct perception of visual phenomena once initial location and feature integration has occurred. The hypotheses summarized here will be addressed by the present research.

5

EXPERIMENT 1.

INTRODUCTION

The present hypothesis is that visual information like object groupings or the edges of objects will both constrain and facilitate feature integration and the perception of objects. If this is the case then there should be different results if the same stimuli are presented with a constraint than when they are presented without the constraint. There are however, three related predictions arising from this hypothesis and all three will be addressed by this experiment. Two predictions are related to the controlling or limiting effects of constraining location information like grouping or the edges of objects, and the third prediction is related to the possible facilitative effects of such constraints.

First, evidence has shown that brief viewing times can lead to illusory conjunctions. If feature integration is constrained by grouping or the edges of objects, then illusory conjunctions should only be formed from features contained in objects that appear within the same group or area. Consequently, a second prediction arises from this first one: that is, illusory conjunctions should not be formed from features contained in two items that fall into different groups or areas of space. I will test both of these predictions in this experiment. Further, if the hypothesized constraints on visual processing actually facilitate correct perception and feature integration, then it follows that when viewing time is brief more efficient processing, and therefore less errors overall, might be expected when those constraints are present than when they are not present. This third prediction will also be tested.

It is suggested that illusory conjunctions arise as an adaptive response to the encoding of degraded information when viewing time is brief (Prinzmetal and Keysar, 1989). However, if illusory conjunctions are formed only from items contained within areas defined by visual constraints, then it follows that they should not occur when no such constraint is present. Treisman and Schmidt (1982) conducted an experiment in which they presented subjects with four coloured shapes after a centrally located attention cue, presented 150ms before the display. This was to test the prediction that a centrally located attention cue would cause the attentional lens to fall onto the shape locations rather than the whole area defined by the digits, and so illusory conjunctions would not be formed with the attentional focus on the central locations. As previously mentioned, it is possible the digit task constituted a subjective boundary which might have led to the letters being grouped and processed as part of one entity or area of space. Also, recall that Nakayama and Mackeben (1989) found that detection of a target was enhanced more if an attention cue was presented between 70ms and 150ms before the target onset, than if the cue was presented concurrently with the target, indicating that attention was captured by the transient, externally controlled attention system. Hence, the precue used by Treisman and Schmidt in this experiment may have constituted a constraining factor that might not only have enhanced detection of the coloured shapes, but also have taken priority over any other external constraining stimuli presented concurrently with the shapes (i.e. the digits). Treisman and Schmidt did not compare the error rates from experiments when the digits were presented with the letters to error rates for this last experiment; they only noted that no significant level of illusory conjunctions occurred in the latter case. In fact they could not have made this comparison as presentation times were varied across their experiments (Tsal, 1989) and to my knowledge, neither has this comparison been made in any of the other research published. The fact that Treisman and Schmidt found no significant level of illusory conjunctions when the primary task was reporting a precued centrally located item, could have resulted because the digits no longer acted as a constraint that could have grouped the letters so they would be processed as one entity or area of space. In fact, in the absence of any more "global constraints", the edges of the items themselves could have acted as "local constraints", thereby leading to their being registered as separate objects or areas, and so to a condition that should prevent illusory conjunctions occurring above chance levels (i.e. illusory conjunctions should not be formed from features contained within different areas of space). I will test the prediction that illusory conjunctions should not occur above chance levels if items are presented in the absence of a constraint.

There is another prediction arising from Prinzmetal and Keysar's (1989) proposal that illusory conjunctions are an adaptive response to the registration of degraded information for some parts of the visual scene when viewing time is brief. They further suggest illusory conjunctions arise because the little visual data that is registered when viewing time is very limited, is spread to fill in the missing data. It is also possible that they could arise from a confusion of the visual information registered if that information was being processed in a parallel system (Townsend, 1990). Either hypothesis leads to the prediction that illusory conjunctions will more readily occur between items that are more similar than between items that are very different. For example, between blue and green more readily than between pink and green. This prediction will also be tested.

An adaption of procedures previously employed with regard to illusory conjunctions is used to test the predictions outlined above. As in Cohen and Ivry's (1989) experiments, subjects are briefly presented with two coloured letters, a target and a distractor, and required to report which of two possible target letters is present in a display, as well as its colour. Following Prinzmetal and Keysar's (1989) experiments there is no primary task, but rather a centrally located fixation point is employed to ensure that the subject's attention is, at least initially, fixed on the centre of the visual display. However, rather than test the effects of grouping on performance as Prinzmetal and Keysar did, I will utilize black square outlines to divide the visual scene into separate areas, and examine the effect these squares have on the occurrence of illusory conjunctions, as well as overall error rates.

Recall that Cohen and Ivry (1989) found illusory conjunctions did not occur above chance levels when the letters were presented 1° or more of visual angle apart. In their second group of experiments, illusory conjunctions were formed when the distance between items was slightly more than 1° of visual angle. Because of their findings the letters in this experiment will be separated by 1° of visual angle. The more conservative distance is used to ensure that the letters will be close enough to lead to illusory conjunctions. The experiment also employs four conditions; one in which only the two coloured letters are presented, one in which both coloured letters are presented inside of the square, one in which both letters are outside of the square, and one in which one letter appears inside and one outside of the square. These four conditions are used to test the predictions.

1. First, it is predicted that illusory conjunctions will only be formed from features contained in objects that have been grouped together in some way. This prediction will be tested by presenting the two letters so they both appear either inside or outside of the square. If the square acts as a constraint and groups the letters, resulting in their being processed in parallel as part of a single area or entity, then illusory conjunctions should occur above chance levels. The square should also lead to the two letters being treated as belonging to a separate single area or entity when they are outside of the square. Therefore, illusory conjunctions should occur above chance levels when the two letters both appear either inside or outside of the square.

2. It is also predicted that constraints on feature integration should prevent illusory conjunctions from being formed from features contained in two items when those items each fall into different groups or areas defined by those constraints. To test this prediction, the two coloured letters will be presented on some trials so that one letter falls inside the square and the other falls outside the square. It is expected that illusory conjunctions will not occur above chance in this condition, and further, that they should occur at a rate which is less than would be expected by chance.

3. It is predicted that constraints on visual processing should facilitate the correct perception of stimuli. To test this prediction total error rates are compared between a condition in which only two coloured letters are presented, and conditions in which the letters are presented with a black square. If the square acts as a constraint on feature integration by grouping the letters, resulting in their being processed in parallel as a part of the area into which they fall, then more accurate target detection should occur when the letters are presented with the square than when they are presented without the square.

4. It is predicted that illusory conjunctions should only occur above chance levels when items are contained within an area defined by a constraint and further, when no such constraint is present, illusory conjunctions should not occur above chance levels. To test this prediction illusory conjunction rates for the condition when the letters appear with the black square will be compared to illusory conjunction rates for the condition when the letters are presented with no black square. While illusory conjunctions should occur above chance levels when the letters are grouped together by the square, they should not occur above chance levels when no square is present.

5. Finally, the prediction that illusory conjunctions are more likely to occur when two items are more similar in colour than when they are less similar, will also be tested: for example, they might more readily occur between blue and green items than between blue and pink items. To test this prediction, illusory conjunction rates are compared across the six colour pairs derived from the four colours used with the letters in this experiment.

And now, it is time "to leave the prim path of rosy speculation and muck about with the data" (Ziff, 1964, p214).

METHOD

Subjects.

Twenty four subjects participated in the experiment. All were secondary school and university students aged between 17 and 43 years of age (Appendix A, 1, Table 1). Eighteen of the subjects were females and six were males. All of the subjects informed the experimenter that they had normal or corrected to normal vision.

Apparatus and Stimuli.

Stimuli were presented with slides using three tachistoscope projectors which were controlled by an Apple SE computer, using software programmed in Hypertalk, via an IDAC controller programmed in Pascal. All stimuli were projected from a projector room through a hatch in the wall onto the back of a rectangular screen surrounded by a black cardboard frame. This frame measured 40.2 cm in height by 50.7 cm in width and the rectangular screen hole cut in this frame was 16cm in height by 24 cm in width. The screen was situated in the experimental room lit by one Superlux desk lamp with a 100 watt bulb. This lamp was placed behind the subjects' chair and adjusted so that it shone towards the upper part of the wall on which the screen was situated (Appendix A, 2).

Figure 13a Figure 13b Figure 13c

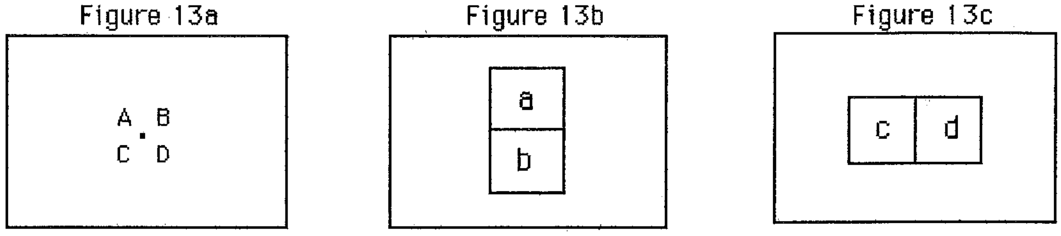


Figure 13a shows the central fixation point (not present during the letter display) and the 4 letter-pair locations: an upper horizontal position (AB), a lower horizontal position (CD), a left vertical position (AC), and a right vertical position (BD).
Figures 13b and 13c show the 4 black square outline locations: an upper (a) and a lower (b) square (Fig. 1b), and a left (c) and right (d) square (Fig. 1c). All these squares were placed so that the center point of the inner edge fell on the central _fixation point of the display.

The stimuli contained in the slides consisted of an asterisk, the letters O, S, F and V which were coloured pink, yellow, blue or green, black square-outlines, and multi-coloured geometric shapes. There were 290 slides. One slide contained only a centrally located fixation point in the form of small black asterisk. A second slide was used as a mask and consisted of a variety of multi-coloured geometric shapes. The colours of these shapes differed from the four colours assigned to the letters. The remaining 288 slides each contained two coloured letters. One of these letters was always an F or a V (the target letters) while the other letter was an O or an S. Half (144) of the slides contained only the two letters, and on each slide these were assigned to one of four possible letter-pair locations. The letter-pairs appeared in each of the four locations equally often. Figure 13a shows the four letter-pair positions and how the two letters on each slide could be located horizontally in either the upper or lower positions, or vertically in either the left or right positions. All four letters were an equal distance from the central fixation point.

Figure 14a

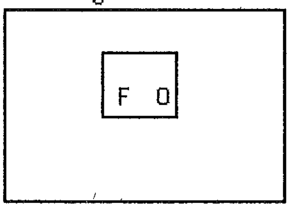


Figure 14b

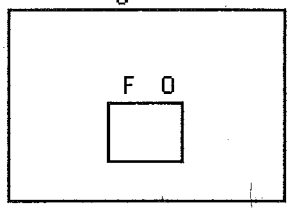
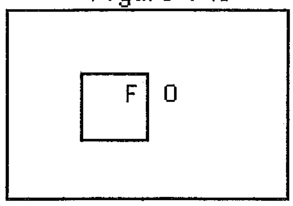


Figure 14c



<u>Figures 14a, 14b, and 14c</u> show the upper horizontal letter-pair (FO) in the 3 shape conditions. In Fig. 14a the "in-shape" condition is shown with both letters inside the upper square (a in Fig. 13b). In Fig. 14b the "out-shape" condition is shown with both letters outside the lower square (b in Fig. 13b). In Fig. 14c the "divided-shape" condition is shown with one letter (F) inside, and one letter (O) outside the left square (c in Fig. 13c). The horizontal letter-pairs were also combined with a right square (d in Fig. 13c) in order to create displays in the divided-shape condition.

The remaining 144 slides were exact replications of the first 144, but with the addition of a single black square-outline that appeared in each of four possible locations in equal numbers of slides. Figures 13b and 13c show these four possible locations. The squares were arranged so that the the two letters in each letter-pair location appeared equally often in each of three possible shape conditions: both letters inside the square (the "in-shape" condition); both letters outside the square (the "out-shape" condition); or with one letter inside and one outside the square (the "divided-shape" condition). Figures 14a, 14b and 14c show the 3 shape conditions for an upper horizontal letter-pair. The 144 slides that contained only coloured letters effectively created a fourth condition (the "no-shape" condition) (Appendix A, 3).

The letter pairs were arranged so that the target and distractor letters appeared equally often in the various possible combinations. Each of the target letters was combined equally often with each of the two distractor letters, and the order of the letters was reversed for half of the presentations in each of the four letter-pair locations (e.g. target letter on the left and distractor on the right for half of the presentations in the horizontal locations, and the order reversed for the other half of the presentations). Also, colours were not randomly assigned to letters, but were assigned so that each letter pair appeared equally often in each of the six possible colour pairs (i.e. blue-green, blue-yellow, blue-pink, green-yellow, green-pink, and yellow-pink). Colour pairs also appeared equally often in both orders (i.e. yellow-pink and pink-yellow, for example). The squares were assigned to letter pairs so that colour and letter pair combinations appeared equally often in each of the three shape conditions. The 288 slides were randomly assigned to six blocks of forty eight slides each. Copies of a selection of these 288 slides were used to create six practice blocks of twenty four slides each.

Subjects were seated so the distance between their eyes and the screen was 180 centimeters and the distance between the two letters was approximately 1° of visual angle. Visual angles as well as letter and square dimensions are summarized in Table 1 (Appendix A, 4).

TABLE 1: The screen surface distances (w) and visual angles (Ø) for Experiment 1, where the subjects' distance from the screen (d) = 180 cms (figures may be approximate because of rounding).

	w (in cms)	Ø
Between letters	3.1	1.0°
Letter:		
height	1.4	0.44°
width	1.1	0.33°
Square sides	6.5	2.1°
Screen:		
height	16.0	5.1°
width	24.0	7.5°

Procedure.

The experiment consisted of six blocks of forty eight trials each, preceded by between three and six practice blocks of twenty four trials each. Possible effects of fatigue or practice were controlled for by changing the presentation order of the experimental blocks in a systematic way across subjects. At the beginning of each trial a central fixation point, in the form of a small asterisk, was presented. Subjects were instructed to keep their gaze and attention fixed on this point. After 1000 ms the fixation point was replaced by a display containing two coloured letters in one of the four possible letter-pair locations described. This display was subsequently replaced by the presentation of a mask for 1000 ms, and in the form of multicoloured geometric shapes. Subjects were told that a target letter, in the form of either an F or a V, would be present in every trial. They were instructed to respond verbally by saying F or V, whichever letter they had perceived, and then to state its colour using one of the four possible colours used in the experiment. In half of the trials this display also contained a black square outline but participants were informed that no response was required in relation to these squares. Subjects were also instructed to guess in cases where they had missed seeing the target letter or were unsure of their perceptions. The subjects' responses were entered, by the experimenter, into the Apple SE computer situated in an adjacent room. On completion of each block the colour, letter, illusory conjunction, and total error rates for all four conditions for that block were calculated by the computer.

During the first practice block the letter display was presented for 200 ms. The presentation time was reduced to 150 ms, 100 ms, and then by decrements of 16.6 ms (1/60 sec) on subsequent practice blocks, until each subject reached an error rate of about twenty five percent. If the subject's error rate was still as low as ten percent or less at the end of the third practice block, then a decrement of 33.3 ms (2/60 sec) was made in the presentation time for the next practice block and, if required, further decrements of 16.6 ms were made on subsequent blocks. After completing between three and six practice blocks, each subject completed the six experimental blocks. If the error rate changed substantially during the experimental blocks (i.e. fell below 15% or increased to over 35% on two consecutive blocks) the letter display time was adjusted by 16.6 ms (Appendix A, 5). Particularly at the faster presentation times, decrements of 16.6 ms could push a subject's error rate from under fifteen to over thirty five percent, but technical limitations prevented smaller time adjustments being made. In these cases the presentation time producing the higher error rate was maintained resulting in some subjects' total error rates being substantially higher than twenty five percent. For other subjects, two consecutive blocks at less than a fifteen percent error rate led to an overall rate of substantially below twenty five percent. During the experimental blocks no feedback was given with regard to performance.

RESULTS

The mean projection time for the letter displays for this experiment was 5/60ths of a second and ranged from 2/60ths to 6/60ths of a second (Appendix B, 1, Table 1). The mean projection time was similar to times used in previous related research (e.g. Cohen &Ivry, 1989; Prinzmetal & Keysar, 1989; Treisman & Schmidt, 1982). The average total rate of errors was twenty eight percent. Five error types are possible with the task employed in this experiment. Feature errors occur when one feature of an item is reported correctly and one reported incorrectly. There are two types of feature error. *Letter feature errors* arise where the colour of the target is reported correctly but an incorrect letter is given. *Colour feature errors* occur where the target letter is correctly reported but the colour reported is one not present in the display. *Illusory conjunctions* occur when two features, each from a different item in a display, are reported as an item that is not actually present. In this experiment the subjects were informed that the letter was always an F or a V and this effectively prevented the report of illusory conjunction errors that combined the distractor letter with the other colour in the display. Illusory conjunctions were therefore indicated by a correctly reported target letter but as having the colour of the distractor (i.e. the other colour present in the display). *Complete errors* occur when both the target letter and its colour are reported incorrectly, so both a letter feature and a colour feature error are made for the same target. *Letter feature plus conjunction errors* are complete errors, but here an incorrect report of the target letter is combined with the colour of the distractor, as it is for illusory conjunctions. However, these errors do not constitute illusory conjunctions because only one of the reported features is present in the display (Appendix B, 6, Tables 1 to 5 give raw data for the five error types).

A critical analysis for this research is of the rate of illusory conjunctions in each of the four conditions. Following the method used by Cohen and Ivry (1989), chance rates of illusory conjunctions were considered as a ratio of 2:1 colour feature errors to illusory conjunctions. Four colours were used and on any given trial one of these four colours would be the correct colour, leaving three possible error colours: one present in the display in the distractor, the mention of which would constitute an illusory conjunction; and two not present in the display, either of which would therefore constitute a colour feature error if reported. One-tailed *t-tests* were used to test for rates of illusory conjunctions occurring significantly above in the no-shape, in-shape and out-shape conditions and below chance in the divided shape condition. If a ratio of 2:1 colour feature errors to illusory conjunctions is the chance level of illusory conjunctions, then when illusory conjunctions occur above or below half the rate of colour feature errors, they would be occurring above or below chance. The *t-tests* were therefore performed using the mean of the differences in rates of colour conjunctions and half the colour feature errors, following Cohen and Ivry (1989) (Appendix B, 2).

The *t-tests* showed that, as predicted, illusory conjunctions occurred at a rate significantly above chance for both the in-shape condition, $t(23) = 3.00$, $p < 0.005$, *1-tailed*, and out-shape condition, $t(23) = 6.22$, $p < 0.0005$, *1-tailed*. While illusory conjunctions did not differ significantly from chance levels for the no-shape condition, they occurred at a rate significantly below chance for the divided-shape condition, $t(23) = -4.89$, $p < 0.0005$, *1-tailed*, (Appendix B, 2, Table 1). Table 2 shows the mean rates of the different types of errors made by subjects, as proportions of the total responses made in each of the four conditions. Illusory conjunction rates significantly above or below chance levels are marked by an asterisk.

TABLE 2: Mean proportions of the different types of error responses made in four conditions for Experiment 1, where the visual angle between letters was 1°.

	Conditions			
Type of response	No-shape	In-shape	Out-shape	Divided-shape
Colour feature	.139	.121	.098	.161
Illusory conjunctions	.065	.087*	.096*	.011*
Letter feature	.102	.05	.069	.096
Complete error	.012	.005	.004	.008
Letter feature plus colour conj.	.004	.002	.002	.002
Total errors	.322	.265	.269	.278

Results of the *t-tests* were as predicted and appear to indicate that the inclusion of the black square, and its relationship to the two coloured letters, are the critical factors in determining whether illusory conjunctions occur at a significant level when the letters are separated by 1° of visual angle. When one letter was placed inside the square and one outside the square, illusory conjunctions occurred at a rate much lower than would be expected by chance. Also, while illusory conjunctions occurred at an above chance level when both letters were either inside or outside the square, they occurred only at a chance level when no square, and therefore no constraint, was present.

As well as using *t-tests* to test for a significant level of illusory conjunctions within each condition, the means of the differences in rates between illusory conjunctions and half of the colour feature errors, were compared across the four conditions using an analysis of variance. This analysis revealed a significant conditions effect, $F(3,23) = 45.27$, $p < 0.00001$ (Appendix B, 3, Tables 1 and 2). Tukey tests indicated that the proportions of illusory conjunctions in the in-shape and out-shape conditions were not significantly different. Significantly more illusory conjunctions occurred in both these conditions than in the no-shape and divided-shape conditions. A significant difference in illusory conjunction rates was also found between the no-shape and divided-shape conditions. As predicted, the divided-shape condition elicited significantly fewer illusory conjunctions than all of the other three conditions (Appendix B, 3, Table 3).

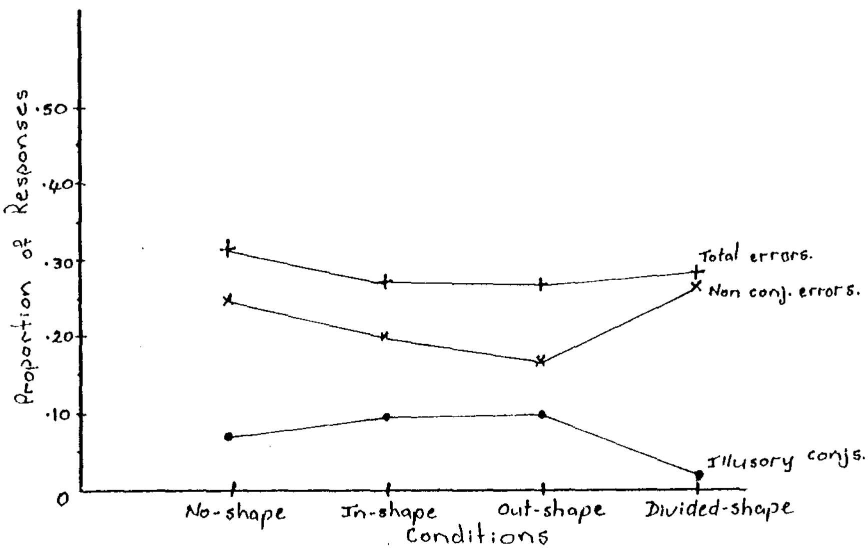
Together with the t-test results, these results show that illusory conjunctions occurred significantly above chance levels only when both letters appeared either inside or outside of a square. While illusory conjunction rates did not differ from chance levels when the letters were presented without a square, they occurred at levels significantly below chance when one letter was inside and one outside of the square.

The other important analysis for this research is of the overall, or total error rates elicited by each of the conditions. It was predicted that the black square should facilitate more accurate detection of the target, and therefore lower error rates, in the shape conditions than in the no-shape condition. An analysis of variance comparing the mean total error rates across the four conditions revealed a significant effect, $F(3,23) = 7.19$, $p < 0.0003$, (Appendix B, 4, Tables 1 and 2). Tukey tests indicated, as predicted, that this effect could be accounted for by the no-shape condition eliciting significantly more errors than any of the shape conditions, none of which differed significantly from each other in total error rates (Appendix B, 4, Table 3).

This predicted result seems to indicate that the presence of a constraint in the form of the square facilitated better detection of the target than when no constraint was present. However, a close scrutiny of Table 2 suggests that comparing the total error rates may not be totally informative. The data in table 2 indicates the divided-shape condition elicited feature error rates that are substantially higher than those for either of the other shape conditions. In fact it produced feature error rates that are as high, or higher than in the no-shape condition. It would appear then, that the significantly below chance rate of illusory conjunctions elicited by the divided-shape condition, has reduced the total error rate and masked a higher rate of the other types of errors in this condition than in the other shape conditions. Therefore, the total errors minus the illusory conjunctions (i.e. the non conjunction errors) were examined. An analysis of variance of non conjunction errors revealed a significant effect, $F(3,23) = 19.745$, $p <$

0.00001, (Appendix B, 5, Tables 1 and 2). Tukey tests indicated that more non conjunction errors were made in the divided-shape condition than in the other two shape conditions, and that these errors were equal in frequency in the divided-shape and no-shape conditions (Appendix B, 5, Table 3). Figure 15 demonstrates the differences in illusory conjunction, non conjunction and total error rates across the four conditions.

Figure 15: Mean illusory conjunction, non conjunction and total error rates, as proportions of the total responses for each of 4 conditions in Experiment 1 (n = 24).



As predicted, error rates were lower for the in-shape and out-shape conditions, even though illusory conjunction rates were higher for these conditions than the other two conditions. However, in the divided-shape condition (when one letter was inside and one outside the square) the illusory conjunction rate was significantly lower than for any of the other conditions, but an analysis of non conjunction errors indicates that these errors were significantly more likely in the divided-shape condition than for the other shape conditions, and in fact more similar to the no-shape condition. So, when the lower rate of illusory conjunctions for the divided-shape condition is taken into account, error rates for this condition appear to be more similar to the no-shape condition than to the other shape conditions.

Finally, recall that it was predicted that illusory conjunctions would be more likely to arise when more similar colours were presented together, than when colours that are very different were presented together (e.g. blue and green rather than say blue and pink). Table 3 (p87) presents the relevant raw data for the twenty four subjects. While the nature of the data precluded the use of parametric statistical procedures, it is clear illusory conjunctions are more likely for some colour pairs than others. For example, it can be seen that all twenty four subjects produced illusory conjunctions for the blue/green pair, but not for any other colour pair (a mean of 0.332 compared to 0.005, 0.002, 0.006, 0.003, and 0.041). So, as predicted, illusory conjunctions appear more likely when the two letters presented are similar in colour (i.e. blue and green) than when they are very different in colour (e.g. blue and yellow).

TABLE 3: Illusory conjunction rates as proportions of each subject's total responses for each of 6 colour pairs, as well as means and standard deviations for each colour pair in Experiment 1 (n = 24).

	Colour pair					
Subject	Blue/green	Blue/pink	Blue/yellow	Green/pink	Green/yellow	Pink/yellow
1	.292	.021	0	0	0	0
2	.438	0	0	0	0	.021
3	.333	0	.021	0	0	0
4	.188	0	0	0	0	0
5	.292	0	0	0	0	.083
6	.438	0	0	0	0	0
7	.125	0	0	0	0	0
8	.292	0	0	0	0	.083
9	.25	.021	0	0	0	.021
10	.313	0	0	0	.021	0
11	.208	0	0	0	0	0
12	.458	0	0	0	0	.146
13	.417	0	.021	.021	0	.104
14	.333	0	0	0	.021	.104
15	.313	0	0	0	0	0
16	.375	.021	0	.021	0	.083
17	.5	0	0	0	0	.063
18	.354	0	0	0	0	.021
19	.229	0	0	0	.021	.021
20	.479	0	0	0	0	0
21	.479	0	0	0	0	.083
22	.354	.021	0	0	0	.042
23	.208	0	0	0	0	.083
24	.292	.042	0	.104	0	.021
Means	.332	.005	.002	.006	.003	.041
Stdevs.	.102	.011	.006	.022	.007	.045

DISCUSSION

The results from Experiment 1 demonstrate that constraining visual information is a critical factor in determining whether illusory conjunctions are formed or not. When both letters were either inside the area defined by the square outline or both in the area outside of the square, illusory conjunctions occurred at a level that far exceeded chance. However, when the square was not present, illusory conjunctions did not occur above chance levels. Further, the results demonstrate that constraints actually prevent illusory conjunctions being formed from features contained in items or objects that fall into different areas defined by such constraints. When one item fell inside and one outside of the square, the colour of one letter was prevented from spreading to, or becoming confused with, the colour of the other letter. So, illusory conjunctions not only arise as a result of the presence of a constraint, but are also prevented by the presence of such a constraint.

The results from Experiment 1 indicated there was another factor involved in determining whether illusory conjunctions would be formed or not. This was the similarity of features found in separate objects. Illusory conjunctions were elicited far more frequently when items were similar in colour than when they were coloured very differently. When the two letters were coloured blue and green far more illusory conjunctions occurred than when they were coloured blue and yellow or green and pink, for example. However, while the blue and green appeared to be more similar in colour than any other colour pair, due to technical limitations I could not ascertain in what way hue, saturation, brightness or a combination of these dimensions is responsible for the colour similarity effect. Nevertheless it is possible that at least some of the large differences apparent in Table 3 could be due to colour spreading when brief viewing times led to degraded visual information, or even a colour additive effect as Prinzmetal and Keysar (1989) suggest might be the case.

The proposal of a colour additive effect rests on the idea that the colours of two items might be combined to fill in the missing data. For example, red and green might be combined to form brown. Some anecdotal evidence supports the notion that illusory conjunctions involving colour arise from a colour additive effect, as well as indicating how the colour similarity factor might be better tested. Several subjects reported seeing the target letter in a colour that was different to any of the four colours used. An example of one such colour were reports of purple and, after several such reports, a check on the actual colours presented when subsequent reports of purple were made revealed those colours to be pink and blue. Perhaps allowing subjects the choice of more colour responses (eight for example) might produce results that indicate a colour additive effect. It would also be useful to acquire similarity ratings of the actual colours used in order to obtain an initial colour-similarity measure. Table 3 certainly indicates that the similarity factor warrants further exploration.

It was predicted that less errors overall should occur when the square was present than when it was not present. This prediction was derived from the hypothesis that constraining visual information might facilitate the accurate detection of objects by dividing visual space into units or chunks, and the information within those units would then be processed together in a limited capacity parallel system.[1] Besides demonstrating that the occurrence of illusory conjunctions is determined by more than one factor, the results of Experiment 1 supports the hypothesis that constraining visual information might facilitate the accurate detection of objects. The results demonstrate that the black square had considerable effect on the ability to correctly identify the target letters. Further, these results show that not only the overall error rate, but also rates of all of the different types of errors are determined by at least two factors: the presence of a constraint, and the relationship between the items and the constraint.

First, as predicted, the presence of a constraint not only leads to illusory conjunctions, but also facilitates accurate detection of objects. That is, more items were correctly detected when they both fell either inside or outside of the constraining square than when the square was not presented. Second, an analysis of non conjunction errors showed that the pattern of these errors differed from the pattern of total error rates, and the pattern of illusory conjunction rates, across the four conditions. That is, where the letters were situated in relation to the square determined not only the rate of errors but also which type of error was more likely to occur. When the square was not present, in the no-shape condition, illusory conjunctions occurred at chance level but more non conjunction errors occurred than

[1] I have proposed that such processing would occur in a limited capacity parallel system because the confusion or combination of features from different items, which occurs in the formation of illusory conjunctions, would be more likely to happen if the items were processed together in a parallel system than if they were processed separately, and probably sequentially, in a serial system (Townsend, 1990).

the in-shape and out-shape conditions. However, when both letters appeared either inside or outside of the square, illusory conjunctions were above chance levels but non conjunction errors were far lower than for the other conditions, resulting in an overall error rate that was significantly lower than for the no-shape condition. In contrast, when one letter appeared inside and one outside the square, non conjunction errors were as high as for the no-shape condition, but an illusory conjunction rate far below chance level led to a total error rate similar to that elicited by the other two shape conditions (Figure 15 in the results section, demonstrates the way each type of error differed across the 4 conditions)

Although the pattern of non conjunction errors across the four conditions was not predicted, the results still support the prediction that the black square would lead to more accurate detection of the target than would occur when no square was present. This is particularly apparent for the in-shape and out-shape conditions where total error rates were lower than for the no-shape condition, despite the fact that these two shape conditions elicited much higher rates of illusory conjunctions than the other conditions. The prediction of fewer errors occurring when the square was present than for the no-shape condition is true only of illusory conjunction rates for the divided-shape condition. Non conjunction rates were not less frequent in the divided-shape condition than in the no-shape condition. This result will be discussed further when it is compared to the results from experiment 2.

In Experiment 2 the visual angle between the letters will be increased to 2° and this is expected to reduce the rate of illusory conjunctions in the in-shape and out-shape conditions. However, the facilitative affects on the accurate detection of the target should still occur. Therefore the results from Experiment 1 with regard to non conjunction errors should be replicated in Experiment 2.

6
EXPERIMENT 2.

INTRODUCTION

Recall Cohen and Ivry (1989) found, that when two letters were separated by more than 1° of visual angle, illusory conjunctions did not occur above chance levels, and further that they occurred at less than chance levels when the letters were separated by more than 2.17° of visual angle. Their results could be explained by the notion that proximity and density acted as constraints on feature integration. That is, when items are separated by more than a critical visual angle, those items are then treated as separate objects or separate areas of space rather than as part of the same object or area, as appears to be the case when they are separated by visual angles of 1° or less.

It might also be expected that the proposed constraints would operate within areas already, or initially, separated by those visual constraints. For example, it would not be adaptive to have the features defining the windows and doors of a building spreading to fill in missing data when brief viewing times or poor illumination led to degraded visual information. It is known that proximity and density appear to constitute constraints on visual processing. Also, Cohen and Ivry (1989) have shown that illusory conjunctions do not occur above chance levels when items located far apart. We might expect than, that when a small number of well separated objects, although fully within a bounded area, will be registered as separate entities rather than as a surface property. If so, illusory conjunctions would not be likely to occur above chance among such items. This is, in effect, suggesting a hierarchy of processing on a different dimension to the one suggested previously (i.e. that visual space is first encoded as separated areas of space and that further processing like feature integration is constrained and contained within those areas). On this suggested second dimension, more "global constraints" are differentiated from "local constraints". That is, the constraints may operate in a type of "depth structure" so that separate (and "local") objects within an initially (and "globally") defined area might be treated as separate entities. For example, separate trees within a forest viewed from a relatively close proximity might be processed as separate objects contained within the initially encoded forest area. If viewed from a great distance, a forest would probably appear more like a texture and the automatic processes involved in constraining visual input may need to be over-ridden in order to identify individual trees, and perhaps require attention to facilitate this identification. This notion could also explain Cohen and Ivry's (1991) density effects on visual search. They found that search times were slower for conjunction targets when they were grouped closely together than for when they were spread apart. Predictions arising from the above hypothesis will be tested in this experiment.

1. The notion of a hierarchical depth structure is tested, and a replication of Cohen and Ivry's (1989) distance effects sought, by conducting a second experiment, identical in nature to the first one except for the visual angles used. In Experiment 1 the two letters were

separated by 1° of visual angle, while in this experiment they are separated by 2° of visual angle. It is predicted that illusory conjunctions will not occur above chance for letters separated by 2° of visual angle, even when a constraint in the form of the black square outline is present. That is, it is not expected that illusory conjunctions will occur above chance levels in any condition in this experiment.

2. "Global constraints" might still be expected to constrain feature integration even when items contained within areas, defined by those constraints, are too distant from each other (i.e.locally constrained) to produce illusory conjunctions. It is therefore predicted that the black square would prevent illusory conjunctions when one letter is inside and one outside of the square. So illusory conjunctions should still occur below chance levels in the divided-shape condition.

3. If the constraints facilitate target detection by constraining attention as well as allowing items within areas to be processed together as single units or entities, then it would still be expected that target detection might be better when the letters are presented together with a square, even though they are separated by 2° of visual angle. It is therefore predicted that total and non conjunction error rates will show the same pattern as in Experiment 1: that is, non conjunction errors should be lower for the in-shape and out-shape conditions than for the no-shape and divided-shape conditions.

METHOD

Subjects.

Twenty four subjects also participated in this experiment. Once again, all subjects were secondary school and university students and their ages ranged between 17 and 43 years (Appendix A, 1, Table 1). Sixteen of the subjects were females and eight were males. All subjects informed the experimenter that they had normal or corrected to normal vision. An attempt to have the same twenty four subjects participate in both experiments proved unsuccessful, resulting in only ten subjects participating in both experiments.

Apparatus, Stimuli and Procedure.

Apparatus, stimuli and procedural details for this experiment were identical to those in Experiment 1, except for the size of the screen and the visual angles cast by the stimuli. While the screen frame area was the same for both experiments, the rectangular screen hole cut in this frame was smaller for this experiment, being 16cm in height by 24 cm in width compared to 21.5 cm by 30.2 cm in the Experiment 1. Also, in this experiment subjects were seated so their eyes were 120 centimeters from the screen and there was 2° of visual angle between the letters. Letter and square dimensions as well as visual angles are summarized in Table 4 (Appendix A, 3).

TABLE 4: The screen surface distances (w) and visual angles (Ø) for experiment 2, where the subjects' distance from the screen (d) = 120 cms.

	w (in cms)	Ø
Between letters	4.25	2.0°
Letter:		
height	2.0	0.88°
width	1.5	0.66°
Square sides	9.0	4.3°
Screen:		
height	21.5	10.2°
width	30.5	14.3°

It should be noted that it is not just the visual angle between the letters that differs from Experiment 1, but also the visual angles of all aspects of the stimuli. It would have been preferable to change the visual angle separating the letters while holding all other visual angles constant, but this would have involved a completely new set of 290 slides.

RESULTS

The mean projection time for the letter displays for Experiment 2 was shorter than for Experiment 1, being 4/60ths of a second, but the range was the same as for Experiment 1, from 2/60ths to 6/60ths of a second (Appendix B, 1, Table 1). The average total rate of errors was seventeen percent for this experiment. This means that briefer display times in this experiment, where the letters were separated by 2° of visual angle, produced a lower error rate than for Experiment 1 when the display times were longer and letters were separated by only 1° of visual angle. That is, subjects were more accurate in their responses when the letters were spread further apart than when they were placed close together, even though viewing times were, on average, shorter when the letters were placed further apart.

As for Experiment 1, the rate of illusory conjunctions in each of the four conditions is of particular interest. For this experiment it was predicted that illusory conjunctions should not occur above chance levels in any condition, but that they should occur at less than chance levels in the divided-shape condition, where one letter is inside and one outside of the square. Once again, one-tailed *t-tests* were performed using the means of the differences in rates of illusory conjunctions and half of the colour feature errors. These *t-tests* revealed that illusory conjunctions did not occur at a significant level above chance in any condition. Also, as predicted for the divided-shape condition, the conjunction errors, once again, occurred at a rate significantly below a level that would be expected by chance, $t(23) = -5.14$, $p < 0.0005$, *1-tailed*, (Appendix B, 2, Table 1). The error rates for this experiment are shown in Table 5 where, once again, colour conjunction error rates differing significantly from chance levels are marked with an asterisk (Raw data for the five error types from this experiment are in Appendix B, 7, Tables 1 to 5).

TABLE 5: Mean proportions of the different types of error responses made in four conditions for Experiment 2, where the visual angle between letters was 2°.

	Conditions			
Type of response	No-shape	In-shape	Out-shape	Divided-shape
Colour feature	.072	.071	.056	.090
Illusory conjunctions	.034	.027	.034	.009*
Letter feature	.074	.043	.05	.079
Complete error	.013	.005	.008	.007
Letter feature plus colour conj.	.003	.003	0	.006
Total errors	.196	.149	.148	.191

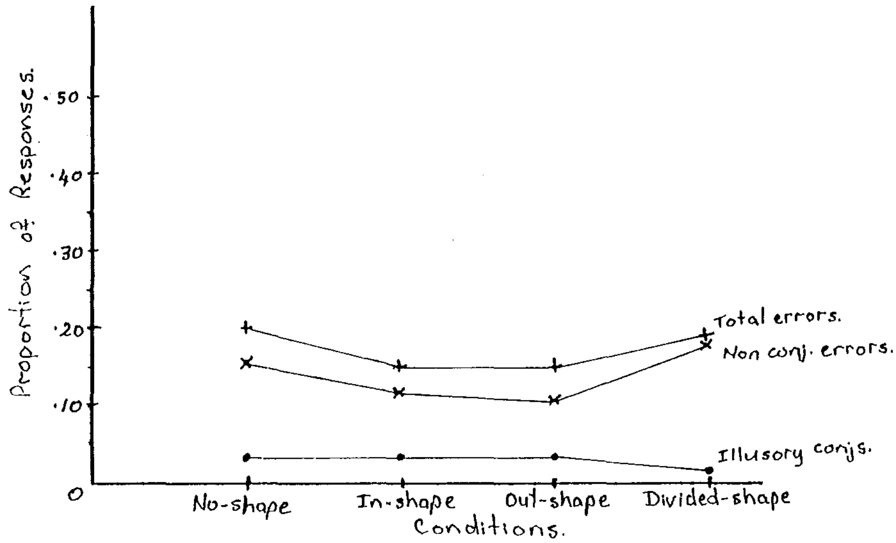
The *t-test* results for Experiment 2 were as predicted in that illusory conjunctions did not occur above chance levels when the letters were separated by 2° of visual angle, even when those letters both appeared in one of the areas defined by the square. When these results are considered together with the results of the *t-tests* from Experiment 1, it can be stated that the distances between objects, as well as constraints in the form of the black squares, affect the rates of illusory conjunctions.

The means of the differences between illusory conjunction and half of the colour feature error rates were compared across the four conditions using an analysis of variance. This analysis revealed a significant conditions effect, $F(3,23) = 9.99$, $p < 0.00001$, (Appendix B, 3, Tables 4 and 5) and Tukey tests revealed, as predicted, that this was accounted for by the divided-shape condition differing significantly from the other three conditions (which did not differ significantly from each other), in having elicited less illusory conjunctions than would be expected by chance (Appendix B, 3, Table 6).

It was predicted that the total as well as the non conjunction error rates should reflect the same pattern that was apparent in Experiment 1. That is, there should be less non conjunction errors in the in-shape and out-shape conditions than in the no-shape or divided-shape conditions. An analysis of variance comparing the mean total error rates across the four conditions revealed a significant effect, $F(3,23) = 10.04$, $p < 0.0001$, (Appendix B, 4, Tables 4 and 5). Tukey tests showed that both the no-shape and the divided-shape conditions elicited significantly greater total error rates than either the in-shape or out-shape conditions, which did not differ significantly from each other (Appendix B, 4, Table 6). This result differs from the pattern of results for total error rates in Experiment 1, where all of the shape conditions produced lower total error rates than the no-shape condition. The result in this experiment better reflects the pattern found when non conjunction errors were analyzed in Experiment 1. Perhaps the lower rate of illusory conjunctions apparent for the in-shape and out-shape conditions in this experiment has resulted in less masking of the higher rate of non conjunction errors, by the very low rate of illusory conjunctions in the divided-shape condition.

In order to evaluate this notion as well as test the prediction that non conjunction errors should occur more in the no-shape and divided-shape conditions than in the in-shape and out-shape conditions, an analysis of variance was used to examine the non conjunction errors. A significant conditions effect was found, $F(3,23) = 14.395$, $p < 0.00001$, (Appendix B, 5, Tables 4 and 5) and Tukey tests indicated the divided-shape condition differed significantly from the two other shape conditions, but not from the no-shape condition (Appendix B, 5, Table 6). Figure 16 demonstrates the differences between illusory conjunction, non conjunction and total error rates for the four conditions in this experiment.

Figure 16: Mean illusory conjunction, non conjunction and total error rates, as proportions of total responses for 4 conditions in Experiment 2 (n = 24).



So, as predicted, subjects were more accurate in their target detection when both letters appeared either inside or outside of the square, than when no square was present or one letter was inside and one outside of the square.

DISCUSSION

As predicted, Experiment 2 demonstrates that proximity is also a factor involved in determining whether or not illusory conjunctions are formed. That is, proximity, as well as the black square and colour similarity shown to effect illusory conjunctions in Experiment 1, are all determinants of illusory conjunction errors. Illusory conjunctions did not occur above chance levels when the letters were 2° of visual angle apart, even when both letters fell into an area defined by the square (i.e. both fell either inside or outside of the square). When this result is compared to those from Experiment 1, it indicates that feature integration is locally constrained (i.e. visual information contained within the square was constrained by proximity) as well as globally constrained (i.e. the visual information was constrained by the black square regardless of the distance between items). Hence foreward I will use the terms "global constraint" and "local constraint" to differentiate between the constraining effects of the square (the global constraint) and the effects of some visual information, contained within an area defined by a global constraint, on illusory conjunction and non conjunction error rates (the local constraints).

In Experiment 1 illusory conjunctions occurred above chance levels when two items were separated by 1° of visual angle, and both fell into one of the areas defined by the square. Also, they did not occur above chance levels in Experiment 2 when the letters were separated by 2° of visual angle, even when both items fell into one area. However, for the divided-shape condition illusory conjunctions occurred below chance levels in both experiments regardless of the distance separating the letters. Further, illusory conjunctions did not occur above chance in the absence of the square in the no-shape condition, regardless of the distance between the letters. This indicates that the global constraint (i.e. the square) continued to constrain visual information even when local constraints (i.e. the proximity of items in Experiment 2) were operating to prevent illusory conjunctions within the areas defined by the presence of the black square.

It is, however, possible the change in the size of the screen, squares, and letters in Experiment 2 also contributed to the results. The visual angle between the letters was adjusted from one experiment to the other by changing the size of the slide projection onto the screen and the subjects' distance from the screen. It might have been preferable to adjust the visual angle between the letters and hold all other distances constant, but this would have required making a whole new set of slides which time limits did not permit. However, these results replicate Cohen and Ivry's (1989) results which indicated that illusory conjunctions did not occur above chance when items were separated by more than 1° of visual angle, and at less than chance levels when items were separated by more than 2.17° of visual angle. Further the change in the size of the letters, in proportion to the increase in acuity with eccentricity from Experiment 1 to Experiment 2, would have prevented visual angle being confounded with visual acuity. But it would be interesting to conduct similar experiments to Experiments 1 and 2, and adjust only the visual angle between items while holding size constant. Perhaps the size of items would interact with the distance between them in determining whether or not illusory conjunctions are formed.

As predicted, the pattern of non conjunction error rates across the four conditions in Experiment 2 replicated the pattern found in Experiment 1 for these errors. However, the results from Experiment 2 show a different pattern in total error rates across the four conditions than occurred in Experiment 1. As we have seen, the pattern of non conjunction rates was the same as for Experiment 1, so the lower rate of illusory conjunctions arising in Experiment 2 than Experiment 1 most likely explains the difference in total error rate patterns. In effect, non conjunction errors occurred at similar rates for the divided-shape and no-shape conditions, and were higher than for the other two shape conditions in both experiments, but the reduction in the rate of illusory conjunctions from Experiment 1 to Experiment 2 in the in-shape and out-shape conditions led to their total error rate being lower than for the divided-shape condition in the second experiment, when it had been similar for all three shape conditions in the first experiment.

Recall that it was originally predicted that the total error rates would be lower for all the shape conditions than the no-shape condition. This has not proved to be the case in Experiment 2 and further, an analysis of non conjunction errors showed these errors occurred at much higher rates in the divided-shape condition than in the other shape conditions, in both experiments. A possible explanation for this result is that the *way* the constraint (i.e. the square) is related to the items not only determines whether or not illusory conjunctions occur but also the rate of other types of errors. While the square may have prevented illusory conjunctions occurring in the divided-shape condition, the way the letters were presented in relation to the square in this condition seems to have increased other types of errors. It is possible that the areas of space defined by the presence of the square were capturing subjects' attention when both letters were placed within one area, thereby facilitating target detection: that is, the "filled" area would indicate in which area the target was located, whether or not that was inside or outside of the square. However, when one letter was inside and one outside the square, both areas would be filled, and there would be no initial indication for the subject of which filled area contained the target. This explanation and the results on which it rests, will be pursued in the general discussion.

7
GENERAL DISCUSSION

The results of Experiments 1 and 2 were as predicted and support the notion of illusory conjunctions being an adaptive response to brief viewing times, in the presence of visual information that constrains feature integration by dividing visual space into objects or areas. The results demonstrate that not only is the presence or absence of such a constraint critical in determining whether or not illusory conjunctions will be formed, but also that there is more than one factor involved in producing illusory conjunctions. Further, the results indicate that the very mechanism which leads to this type of error, also facilitates more accurate detection of objects because it leads to less non conjunction errors than occurs in the absence of such constraints.

In discussing the results of Experiments 1 and 2, it should be kept in mind that the spatial distribution of the letters remained exactly the same whether the square was present or not and irrespective of how the letters were arranged with respect to the square. That is, the letters in a top horizontal letter-pair, for example, did not change in there spatial relationship to each other (in each experiment), whether there was no square present, both fell inside or outside of the square or one fell inside and one outside of the square. It can therefore be stated with some confidence, that the presence of the square and how it was situated with respect to the letters, was instrumental in producing the different effects on error rates across the four conditions. It is also apparent from the results, that the square divided the display into two areas or units (the one bounded by the square and the other outside of the square) because the illusory conjunction, non conjunction and total error rates did not differ significantly between the in-shape and out-shape conditions. It is probable that a if single contour (or black line) divided the display area and was manipulated so it appeared between the letters on some trials and to one side of the letters on other trials, it would yield similar results to those for Experiments 1 and 2.

The results from Experiments 1 and 2 indicate that constraining visual information operates in a kind of "depth structure" so that items falling into an area that has been defined by a constraint (i.e. globally constrained), will also be "locally" constrained so that separate items may be treated as such, even though they may be processed as part of the greater, globally defined, area. For example. a window or door of a house would be treated as a separate part of the house rather than just a surface property. As already mentioned, the results from Experiments 1 and 2 demonstrate that a global constraint not only leads to illusory conjunctions, but also facilitates the detection of objects. Further, even though the distance separating the letters in Experiment 1 resulted in illusory conjunctions occurring above chance levels, the square had the same facilitative effects on the accurate detection of the target in both experiments, when both letters were either inside or outside of the square. Also the results of both experiments indicated that, although subjects had more difficulty detecting the target correctly in the divided-shape condition, they were far less likely to report illusory conjunctions than in the other conditions.

There is a conflict or tension apparent in the claim that a constraint would cause a particular type of error (i.e. illusory conjunctions) and at the same time would also facilitate accurate target detection but, as we have seen, the data actually supports this claim. The claim can be supported, and the pattern of error results from both experiments explained, by a functional approach to visual cognition. First, it would be adaptive to accurately detect the limited amount of visual data that could be registered when viewing time is very brief, especially if that data were used to "fill in" or compensate for the data not registered. Illusory conjunctions appear to be a functionally less serious type of error than feature errors, as they contain only visual information that actually occurs inside an area defined by a constraint. Feature and complete errors provide the viewer with visual data that is not actually present in the object or area. It would probably be less misleading to have information that is contained in an item, even if incorrectly conjoined, than to have information that was not a part of the item. It would therefore be more adaptive when viewing times are very brief, to "fill in" missing data with data registered from the item, rather than with information that did not occur within the area or item in question.

Second, it would also be more adaptive to have non conjunction errors occur than illusory conjunctions, when degraded information is registered from two separate areas or objects, as was the case in the divided shape condition. If features spread, or became confused, from one area to another when viewing conditions were not optimal, visual input might become incomprehensible, and at least be less adaptive in terms of the consequences of acting on the incorrect data, than if feature errors were made. For example, getting the colours of two cars, or a car and the road, confused or mixed could have far more serious consequences than simply getting the colour of a car wrong (i.e. seeing it as orange rather than yellow when the road was grey).

Although a functional approach can explain *why* the illusory conjunction and non conjunction error rates showed the pattern apparent in the results of Experiments 1 and 2, it does not explain *how* such results came about. It is possible that attention played some part in bringing about the pattern of errors apparent in these results, particularly if attention was constrained by the square so that a limited area of the display captured attention (i.e. either the area inside the square or the area outside of the square, but not both). The fact that the non conjunction errors were higher in the divided-shape condition adds support to this notion, particularly if "filled" and "empty" spaces were also differentiated in early processing: a likely event if location and feature information are registered together or in parallel as Johnston and Pashler's (1990) location experiment indicates. This result might be further explained if one also takes account of likely top-down processes that might have been used in completing the task required of subjects in Experiments 1 and 2. Recall it was suggested in the section on visual attention that top-down and bottom-up processes would interact to constrain and capture attention. The task in Experiments 1 and 2 required subjects to find and report a target of which they had prior knowledge. This prior knowledge could have led to fast detection of the target once a constraint divided visual space, and both items within one area were processed in parallel. However, when one letter was presented inside and one outside the square, subjects may have had to override the constraining effects of the square in order to find the target, and perhaps attend to each area serially, rather than both items in parallel as hypothesized for when both letters fall into the same area.

To summarize, the results from both experiments support the prediction that global constraints as well as local constraints determine whether or not illusory conjunctions will be formed. In these experiments colour similarity and the proximity of the letters appeared to be local factors contributing to the formation of illusory conjunctions. However, it is also apparent that a global constraint, such as the black square, is the critical factor in eliciting illusory conjunctions. Illusory conjunctions did not occur in the absence of the square even though the letters were the same colours and distances apart as they were in conditions when the square was present.

Although there was a tension or conflict apparent in the prediction that the black square would both lead to illusory conjunctions and facilitate the accurate detection of items, the results of Experiments 1 and 2 showed that a global constraint could, in fact, facilitate more accurate detection and more illusory conjunctions than when there was no such constraining information (i.e. no square) present. While illusory conjunctions occurred at a higher rate, non conjunction and total error rates were lower in the in-shape and out-shape conditions than in the no-shape and divided-shape conditions. The predicted results therefore support the hypothesis that illusory conjunctions arise as an adaptive response to the registration of degraded visual information, when that information is constrained by visual data which separates space into objects and areas which in turn, facilitates accurate object detection.

The results of both experiments also support the hypothesis that both illusory conjunctions and the facilitative effects of a constraint arise because areas or objects defined by the constraint, are processed in a limited capacity parallel system. It is unlikely that illusory conjunctions would arise if the items from which they were derived were not processed together and in parallel. If features from different items become confused or combined, as is the case with illusory conjunctions, then it is probable they are being processed together, rather than separately and sequentially as they would be in a serial system (Townsend, 1990).

The results of Experiments 1 and 2 will be compared to other related research in the next section. Following that, the implications of these results for theories of visual cognition will be discussed, including Prinzmetal and Keysar's (1989) model, as well as FIT and the Guided Search Model.

OTHER ILLUSORY CONJUNCTION RESEARCH

The results of Experiments 1 and 2 differ from the results of other related research on some points, but also replicate many previous findings. First, as already mentioned, the distance effects apparent in these experiments replicate those of Cohen and Ivry (1989). In their first set of experiments they found that illusory conjunctions occurred above chance levels when items were separated by 1° of visual angle or less, and that they occurred only at chance levels when items were separated by distances between 1° and 2.17° of visual angle. They also found that distances greater than 2.17° of visual angle led to illusory conjunctions occurring at less than chance levels, but the effects of greater distances than 2° between items were not tested in the present experiments. It would be interesting to see how greater distances between the letters would interact with the black square in its effects on illusory conjunctions.

One result obtained by Cohen and Ivry actually conflicts with a result from Experiment 1. Recall that they conducted experiments in which two letters were presented in two of six possible locations. All six locations were in a row, and the experiment utilized a primary digit task. The digits were manipulated to create a "large spotlight" and a "small spotlight" condition. In the large spotlight condition the central four letter locations appeared between the digits, while in the small spotlight condition only the central two letter positions fell between the digits. In the large spotlight condition they found, when the letters were presented so that one fell between the digits and one fell outside of the digits, (i.e. with a letter on either side and adjacent to one of the digits, and the letters were about 1.6° of visual angle apart) that there was a trend for illusory conjunctions occurring above chance level. If, as has already been suggested, the digit task acted as a constraint which grouped the letters, then illusory conjunctions should have occurred below chance levels when the letters appeared on either side of one digit in Cohen and Ivry's experiment. That is, the digits should have acted to prevent illusory conjunctions in the same way as the square outline did in the divided-shape condition in the Experiment 1. However, this difference in results could be explained by Prinzmetal and Keysar's (1989) demonstration that subjective boundaries can act as constraints. When both letters were placed close to, and on either side of one digit in Cohen and Ivry's experiment, this could have led to the whole three items being grouped together and treated as a single area or entity that was separate to the other digit. If this was the case then illusory conjunctions could be expected to occur at greater than chance levels.

Recall that Treisman and Schmidt (1982) found no distance effects on illusory conjunctions in the results of their experiments. Cohen and Ivry's (1989) as well as the present experiments did find distance effects. However, both the present experiments, as well as Cohen and Ivry's, presented subjects with only two coloured letters rather than three letters at once as Treisman and Schmidt did. Treisman and Schmidt presented a row of three coloured letters and found significant levels of illusory conjunctions were formed from features contained in the first and third letters which were separated by more than 1° of visual angle. It is possible that the number of items presented interacts with the distance between those items in producing illusory conjunctions. Perhaps the distance between items that are separated by another item is not relevant as long as the distance between each of the two separated items and the intermediary item is less than 1° of visual angle. That is, as long as a group of items are all in close proximity they will all be incorporated into the same area and processed together. However, if one or more items is separated from the others by a greater distance than separates items in the group, then it should be treated by the visual system as a separate area or object. Texture segregation experiments, that indicate the close grouping of items leads to fast parallel processing, support this notion and it could be tested using an adaption of the method used for the present research. The number or items presented inside the square could be increased and the distances between them manipulated, so they appeared as a single group on some trials and separated into two groups on other trials (e.g. two rows of three items with items within the rows separated by less than 1° of visual angle, and rows separated by more than 2° of visual angle). In fact this method would combine the one used in the present experiments with an adaption of the method utilized by Prinzmetal and Keysar (1989) to test the predictions made by their functional explanation of illusory conjunctions.

The following sub-section will compare the present findings with those of Prinzmetal and Keysar, as well as discuss the implications for their theory.

A functional explanation of illusory conjunctions.

The present results are in agreement with Prinzmetal and Keysar's (1989) experimental results as well as their explanation of illusory conjunctions. Recall that Prinzmetal and Keysar tested the prediction that both subjective and objective groupings would lead to and constrain the formation of illusory conjunctions so they did not arise from combining the features of objects that fell into different groups. Although the present research utilized black square outlines rather than grouping, the results are in agreement with Prinzmetal and Keysar's. Also, like Prinzmetal and Keysar's, the present experiments had no primary digit task when areas of the display were objectively defined, only a central fixation point.

Prinzmetal and Keysar's functional explanation of illusory conjunctions rests on two assumptions. First, that there is poor spatial resolution for some aspects of visual information and second, that spatial information is constrained by perceptual organization. They propose that the constraints are the result of cognitive processes that divide space into separate areas. Further, these processes would reduce spatial uncertainty (i.e. provide some location information) and would constrain feature integration so it did not occur across the boundaries of areas or objects. Further, illusory conjunctions should arise only within areas defined by such constraints, because the constraints would prevent feature integration (or the conjoining of features) across the boundaries of objects or areas. Like Prinzmetal and Keysar's results, the results from Experiments 1 and 2 support this explanation. Illusory conjunctions *only* occurred above chance levels when the two letters were both presented in one area (i.e. either inside or outside of the square). However, the present research went further than Prinzmetal and Keysar's research by demonstrating that the boundaries of areas defined by a constraint actually prevent illusory conjunctions from being formed from features that fall into separate areas. In Experiments 1 and 2 illusory conjunctions occurred at levels less than would be expected by chance in the divided-shape condition. This is clear evidence of illusory conjunctions being prevented by the boundaries of areas or objects. With no constraint present illusory conjunctions would be expected to occur at chance levels; that is at rates no different from any other type of error. This proved to be the case in the no-shape condition in Experiments 1 and 2. If illusory conjunctions had occurred only at chance levels in the divided-shape condition, Prinzmetal and Keysar's hypothesis would not have been supported even though illusory conjunctions were above chance levels in the in-shape and out-shape conditions. It required that illusory conjunctions should be lower in the divided-shape condition than in the no-shape condition to demonstrate that illusory conjunctions would be prevented by a constraint that divided visual space into units or areas. This, of course, was the case in both Experiments 1 and 2.

The current research also tested a prediction derived from Prinzmetal and Keysar's functional explanation. They suggested that their explanation is functionally adaptive in terms of cognitive economy. I therefore inferred that, if such cognitive economy resulted from the visual data within groups or areas being processed together in a limited capacity parallel system, this grouping might facilitate the accurate detection of objects. The results of Experiments 1 and 2 supported this prediction, showing that more accurate detection of the target resulted when both letters were presented inside or outside of the square, than when the letters were presented in the absence of constraining information (i.e. without the square).

The evidence from the current research therefore supports Prinzmetal and Keysar's functional explanation. However, this thesis also extended Prinzmetal and Keysar's functional explanation to include the notion of local constraints. That is, perceptual organizing information like the Gestalt grouping principles (i.e. proximity, similarity, etc.) might serve to locally constrain, or organize, visual information that falls inside an area already defined by a "global constraint". This notion was supported by evidence from Experiments 1 and 2 which showed illusory conjunctions were not formed when items were presented 2° of visual angle apart even when they both fell into the same group or area defined by the square.

To summarize, the results of Experiments 1 and 2 replicate Prinzmetal and Keysar's (1989) findings which demonstrated that illusory conjunctions occur because some visual information actually constrains feature integration by separating visual space into units or areas. The results of Experiments 1 and 2 added further support to Prinzmetal and Keysar's functional explanation of illusory conjunctions by showing that illusory conjunctions are prevented from being formed from the features contained in items that fall into separate areas of the visual scene. Further, both of the present experiments demonstrated that the constraints on visual processing might operate so that visual information such as item similarity and proximity would locally organize visual information that falls inside an area that has been defined by a global constraint.

FEATURE INTEGRATION THEORY

The results of the current research are not easily explained by FIT. According to FIT illusory conjunctions arise either because objects are outside of the boundaries of attention, where features would remain unlocated and might be erroneously conjoined, or because objects fall under the attentional spotlight, but with insufficient time to process individual items serially with focal attention. The present results indicate that even though very brief viewing times lead to illusory conjunctions, the critical factor in determining whether or not illusory conjunctions are formed when viewing times are brief, is the presence or not of a global constraint (i.e. the black square outline) rather than attention. While it is clear that FIT, in its present form, cannot accommodate this result, it is possible this theory of feature integration could explain the results, if it was extended to include an initial stage of processing which divided visual space into areas that constrained feature integration and visual attention. This extension would logically result in a prediction that an initially defined area would capture attention, thereby defining the boundaries of the attentional focus, and then features within the area would be conjoined by attention and serial processing, while features outside of attention would remain unlocated and could therefore be conjoined to form illusory conjunctions. It would also follow, if viewing times were so brief as to prevent serial attention to individual items within the area under the attentional spotlight, that illusory conjunctions would arise inside the attentional boundaries.

There are still major problems for FIT with regard to the results of Experiments 1 and 2, even if it were extended to include constraints in visual processing. First, FIT predicts that illusory conjunctions that combined features from an item inside the boundaries of attention with features from an item falling outside the attentional boundaries, would not occur above chance levels. However, it does not predict illusory conjunctions would be prevented to the extent where they would occur below chance levels, as occurred in Experiments 1 and 2 in the divided-shape condition. Second, the the black squares used in the present experiments were made up of separable features (i.e. four lines of two different orientations), and these would have to be conjoined in order for them to define the separate areas of space that might capture the attentional boundaries. That is, some feature integration would be required before attention selected an area or object, but FIT holds attention as a necessary factor for feature integration. However, it is possible that a square could constitute an "emergent feature". FIT allows for the possibility that detectors for some feature conjunctions might be "hard-wired" into the visual system, so they would be registered in parallel in the same features are proposed to be registered. But according to FIT, features are also initially registered as unlocated or "free-floating" and it is difficult to see how an unlocated square could have constrained the processing of other features. In Experiments 1 and 2 the relationship of the square to the letters was critical in determining whether illusory conjunctions occurred above or below chance levels.

Another major problem for FIT is the different rates of non conjunction errors apparent for the different conditions in Experiments 1 and 2. Even with the addition of an initial constraining stage in visual processing, FIT is unable to explain why non conjunction errors would occur at a much lower rate when both letters were together and either inside or outside of the square, than when the letters were presented without a square, or so that one fell inside and one outside of the square. If illusory conjunctions arise because features remain unlocated and "free floating" through a lack of focal attention then, according to FIT, feature errors would occur in the absence of attention as well as an above chance level of illusory conjunctions. This was the case for the in-shape and out shape conditions in Experiments 1 and 2. However, while FIT predicts the ratio of illusory conjunctions to feature errors to differ for conditions where items are attended or not attended, the theory does not predict fewer non conjunction errors in the absence of attention than when items fall under the attentional spotlight; only more illusory conjunctions if items are not attended. In fact, fewer errors overall should result when items are attended, if attention is the critical factor for feature integration. Therefore, with the addition on an initial stage of processing that divided visual space into areas or units, there are two possible predictions FIT could make about the error rates in Experiments 1 and 2. First, if the letters were close enough to the central fixation point then, according to FIT, the central fixation point might have captured attention in the no-shape condition, so that the letters fell under the attentional spotlight. When the square was present attention would be captured by the square in the way the digits were proposed to in Treisman and Schmidt's (1982) experiments. In this scenario, FIT would predict few errors and a chance level of illusory conjunctions when the black square was not present in the no-shape condition, because the letters would be attended to in the absence of the square. However, a relatively higher error rate and above chance levels of illusory conjunctions would be predicted when both letters fell either inside or outside of the square in Experiment 1. As we have seen, the results of Experiment 1 do not support this prediction.

The other hypothesis that a modified FIT could make about the error rates in Experiment 1, is that the letters were too distant from the fixation point to fall under the attentional spotlight in the absence of the square. If this was the case more errors would be expected than if the letters fell under the attentional spotlight, but illusory conjunctions should also occur above chance levels when there was no square present. Once again, the results of Experiment 1 are contrary to FITs prediction.

It is possible that the differences in non conjunction error rates across the four conditions in Experiments 1 and 2 could be explained by top-down processes. While FIT allows that such processes may be involved in feature integration, by conjoining the features of well learned objects in the absence of attention, it does not specify how these top-down processes might do this. Neither does it specify how such processes might contribute to attentional capture, an important issue raised in the section on visual attention. The Guided Search Model does specify how top-down processes could contribute to attentional selection, and this will be discussed in the next section.

THE GUIDED SEARCH MODEL

Wolfe et al.'s (1989) Guided Search Model does not offer a specific explanation of illusory conjunctions. However, it could explain illusory conjunctions simply by suggesting they arise from feature information contained in separate objects becoming confused or mixed in a parallel system. Recall that the Guided Search Model makes no qualitative distinction between the parallel and serial stages of processing, as FIT does. In fact, the whole process of feature integration described in the Guided Search Model could be accommodated by a limited capacity parallel system. As Townsend (1990) points out, visual data is more likely to be confused in a parallel system than in a serial system which would process items separately and sequentially.

However, in it's present form the Guided Search Model could not account for some of the results of this research. For example, subjects in Experiment 2 produced more correct responses, on average, than those participating in Experiment 1, even though the average display time was shorter in Experiment 2 than Experiment 1. These results are similar to Cohen and Ivry's (1991) density effects on visual search. They found that detection was faster when items were spread apart (as in Experiment 2) than when they were presented closer together (as in Experiment 1). The Guided Search Model cannot account for this result, but as Cohen and Ivry (1991) point out, it could account for density effects if it included a first stage of processing that elicited course location information of some kind.

Like FIT, the Guided Search Model cannot account for the different patterns of illusory conjunction and non conjunction errors apparent in the results of Experiments 1 and 2. No part of the process of feature integration described in the Guided Search Model can explain why non conjunction errors occurred more in the absence of the square and when the letters were separated by the placing of the square, than when both letters were either inside or outside the square. Neither can the model explain why illusory conjunctions occurred above chance in some conditions and below chance in others. However, unlike FIT, the Guided Search Model could account for the pattern of errors apparent in Experiments 1 and 2 if it incorporated an initial stage of processing in which constraints on feature integration and attention divided the visual scene into units or areas.

<u>A functional approach to feature integration and attention.</u>

In this sub-section an explanation of the way illusory conjunctions and non conjunction errors were distributed, across the four conditions employed by Experiments 1 and 2, will be made by incorporating the functional explanation of illusory conjunctions with the Guided Search Model. If the Guided Search Model included an initial stage of visual processing which divided visual space into separate areas and objects, it could account for the results of Experiments 1 and 2 that showed different patterns of illusory conjunctions and non conjunction errors related to whether the square was present or not and to where the square was situated in relation to the letters. That is, the lower rate of non conjunction errors that occurred when both letters fell either inside or outside of the square, than occurred when no square was presented with the letters, or when one letter fell inside and one outside of the square, combined with illusory conjunction rates that were above chance when both letters were either inside or outside the square and below chance when one letter fell inside and one outside the square. As we have seen, this is an aspect of the present results that FIT cannot accommodate, even if it were modified to include an initial stage of processing which constrained feature integration.

The Guided Search Model (see p31 for a more detailed description of this model) proposes that top-down processes arising from expectancies and prior knowledge will contribute to the effective detection of a target in visual search experiments. Wolfe et al. (1989) suggest this happens because such top-down processes will increase the activation of features contained in the target, resulting in increased activation of those features at the target's location once the activation from the features is summed in the Activation Map. This increased activation, in turn, guides attention to the location for which activation is the greatest, and once the item is attended to it can be ascertained if the location does in fact hold the target. The Guided Search Model describes no qualitative difference between the parallel (feature detection) system and the serial system (where features are conjoined), proposing that information from the parallel system is made available to the serial system. So in this model the two systems are not autonomous as is proposed by FIT. In fact, as previously mentioned, the processes described in the Guided Search Model could easily be accommodated by a limited capacity parallel system, which Townsend (1990) suggests could be the most parsimonious explanation of the linearly increasing reaction time curves produced by searches for conjunction targets. Further, the Guided Search Model implies that attention would facilitate object detection, but is not necessary for feature integration. This implication is derived because, in the Guided Search Model, attention is guided to likely object locations by activation levels in the Activation Map, where activation from the feature maps is summed (i.e. activation is summed before attention is guided to potential target areas).

There is another implication that can be drawn from Wolfe et al.'s Guided Search Model. That is, that top-down processes actually constrain or order the allocation of attention. Recall, from the section on visual attention, that both internally controlled, top-down processes and externally controlled, bottom-up processes are likely to interact in constraining attention. As previously mentioned, the results from Experiments 1 and 2, showing that non conjunction error rates differed significantly across the four conditions, can be explained in terms of this interaction. So, if the top-down processes included in the Guided Search Model were combined with an initial stage of processing that divided visual space into areas or units, the non conjunction results could be explained.

Recall that a target detection task, which gave subjects prior knowledge of the two possible targets, was used in the present experiments. When both letters fell either inside or outside of the square, the square could have constrained feature integration so that only features contained within one area were conjoined with each other. Consequently, illusory conjunctions would not result from combining features from objects that fell into different areas or units. It is proposed that the constraints on visual processing are functionally adaptive, in that they lead to economical and efficient use of a limited capacity processing system, and that the visual information contained within an area would be processed in parallel within that limited system, perhaps leading to the confusion of similar features that are being processed together. It is also proposed that the constraints are linked to external environmental cues that provide information which differentiates objects from space and from each other. Further, it is hypothesized that these cues would also constrain attention so whole objects rather than various parts of separate objects and areas are attended to.

In the present experiments when both letters were presented either inside or outside of the square, subjects not only had prior knowledge of the likely target, but also were presented with a filled area (i.e. the one containing the letters) and an empty area. According to the Guided Search Model, the level of activation in the Activation Map would have resulted in subjects attending to the filled area, thereby facilitating the detection of the target. Further, the greater activation for the form features contained in the targets (F or V) because of top-down processes (i.e. prior knowledge), would result in a greater probability of attention being drawn to the target than to the distractor, also increasing the likelihood of accurate target detection. Also, with the square present, its constraining effects should lead to items within a filled area being processed in parallel, resulting not only in accurate target detection but also in illusory conjunctions if viewing times are brief. This was the case in Experiment 1.

When one letter appeared inside and one outside of the square, both areas would have contained a letter, and therefore been registered as filled. If the square constrained visual processing so that only the features contained within one area could be processed together, then the letters would not be processed in parallel as would be the case when both fell into the same area, and illusory conjunctions would be prevented. However, the only indication of which area contained the target would be the slight increase in activation for the target location resulting from prior knowledge of the target letters. This is in contrast to when both letters fell into one area, when all activation would come from only one "filled" area or unit. So, more non conjunction errors as well as less illusory conjunctions would be expected when each letter fell into a different area than when they both fell into the same area. Further, if attention was also constrained by the square, then it is likely that the two areas would have to be attended to separately and sequentially in the divided-shape condition. Very brief viewing times would, of course, make it unlikely that both items would be attended to. The inability to attend to both items should further increase the probability of, not just less illusory conjunctions, but also of more non conjunction errors than would occur when both letters fell into the same area. This, of course, was the pattern of results in Experiments 1 and 2.

It is also apparent from the error rate elicited by the divided-shape condition, that the selection of the potential target area was not entirely random. The non conjunction error rate, although being higher than in the other shape conditions, was not high enough to indicate a random choice or guessing. As has already been pointed out, the square itself would provide no information as to which area contained the target. It would provide only cues that effectively prevented the two letters from being processed together. However, the increased activation the Guided Search Model suggests occurs because of prior knowledge of the target, should lead to accurate detection of the target some of the time which did happen in the divided-shape condition.

The explanation just described is further supported by the fact that the non conjunction error rate was as high when there was no square presented with the letters, as it was for when one letter was inside and one outside the square. Once again, prior knowledge would lead to increased activation of the target area. However, with no constraint present that would lead to the parallel processing of the items as if they belonged to one area or unit, there should be no significant level of illusory conjunctions. Further, the two letters should be attended separately and sequentially, as would be the case when one letter appeared inside and one outside of the square. Once again, very brief viewing times would make it unlikely that both items would be attended to in these experiments.

In summary, it was not originally predicted that higher non conjunction error rates would arise in the divided-shape condition than in the other shape conditions in these experiments. However, this result can be accommodated by combining an extended version of Prinzmetal and Keysar's (1989) functional explanation of illusory conjunctions with Wolfe et al.'s Guided Search Model, to create a more comprehensive and functional theory of visual cognition, that includes top-down as well as bottom-up processes.

CONCLUSIONS

Psychology has long been concerned with finding the basic principles underlying human thought and behaviour (Allport, 1989). Medin and Wattenmaker (1987) suggest that mechanisms originally hard-wired into basic processes like visual perception may, in human evolution, have gradually become accessible to the "higher order" cognitive processes that gave rise to consciousness and flexibility. They further suggest that by unravelling the constraints on perceptual and basic cognitive processes, we may also discover the principles that govern or constrain higher order processes. Dividing perceptual information into "chunks", groups, or units appears to be a domain general phenomenon, that occurs at the perceptual level as well as in higher-order or more abstract levels of processing. Evidence has shown it is a factor in speech and language, categorization, conceptual structures, music, reasoning, and memory. In all of these domains, dividing perceptual information into groups or units contributes to cognitive economy (Roche, 1978), as well as accuracy (Miller, 1956) and flexibility (Neisser, 1987). It is hardly surprising then, that current evidence is suggesting that visual cognitive processes like feature integration, as well as visual attention, may be constrained by perceptual grouping. However, defining what those principles are is an important issue.

If perceptual grouping arose early in human evolution, it is probable that such grouping would be linked to the cues in the environment, because perceptions are of the environment and also, in our behaviour we interact with the environment. Also, there is evidence to suggest that we are "hard-wired" to perceive boundaries at the perceptual level in audition (Flavell, 1985), and evidence from the present research as well as from others (Cohen & Ivry, 1989; Prinzmetal & Keysar, 1989) suggests this is also the case in visual perception. The present research suggests that the edges of objects would act as constraints, and together with Prinzmetal and Keysar's (1989) as well as Cohen and

Ivry's (1989) research suggests the Gestalt grouping principles like proximity and similarity may also constitute constraints on visual cognition. Because texture segregation obviously divides visual space into areas, it is also likely that Julesz' (1975, 1981, 1984) textons, in conjunction with the Gestalt grouping principles, may serve to constrain visual processing. Textures, like colours, are surface properties and would contribute to defining the boundaries of objects or areas. Brightness contrast can also define the boundaries of objects and areas. The other obvious candidates for basic constraints on visual processing are the phenomena that have been shown to automatically capture peoples attention. These include movement and flicker (Johnston et al., 1990; Posner et al., 1980). Also, phenomena that have led to "pop-out" in visual search experiments, when targets have contained a conjunction of the features present in the distractors, are also strong candidates as constraints on visual processing. These include binocular disparity (Nakayama & Silverman, 1986), direction of lighting, that might act as a depth cue (Enns & Rensink, 1990) and movement parallax (McLeod, Driver, Dienes & Crisp, 1991).

It is also apparent from Prinzmetal and Keysar's (1989) research, that subjective as well as objective grouping can constrain visual processing. Recall that they presented subjects with an evenly spaced matrix of items and found illusory conjunctions only arose from features contained within a row when subjects were required to complete a primary digit task with the digits presented on either side of the matrix. Similarly, illusory conjunctions were only derived from features contained within a column when the digits were placed above and below the matrix. This result implies that we may be hard-wired to divide visual space into areas even in the absence of any obvious external constraint like proximity, and further, that there is a certain amount of flexibility in what might constitute a constraint. In the case of Prinzmetal and Keysar's results it seems that the grouping was *inferred* from the presence and placement of the digits.

Finally, two areas of conflicting evidence may be resolved by a functional approach to visual cognition and attention. Although not a central issue for this thesis, the first is the unresolved debate over whether visual attention is object or spatially based and whether attentional selection occurs early or late in visual processing. That is, it is suggested by these different approaches to attentional selection that attentional selection would be spatially based if it occurred early in visual processing, before features were combined into objects, or objects were recognized, or it would occur late in visual processing after feature integration and object recognition had occurred. Allport (1989) points out that organisms interact with objects and at first, a functional approach to this debate would seem to suggest that attention would be object based. However, it is apparent from the present research as well as in Prinzmetal and Keysar's (1989) research, that constraints on visual cognition and attention operate early in visual processing. This implies that Prinzmetal and Keysar's assumption that attention is spatially based is correct. Given that the constraints appear to act so that visual space is divided into objects and areas, spatially-based early attentional selection is not incompatible with the notion that it would be adaptive for attention to select the objects in the environment with which organisms interact. However, it has also been demonstrated that top-down processes may constrain visual processing and attention, and that the propensity to divide space into units or areas may itself be hard-wired allowing flexibility as to what might act as constraints. Prinzmetal and Keysar's (1989) results indicated that illusory conjunctions were contained within groups that had been subjectively defined. Their findings suggest that the groupings were inferred from the way the digits were placed in relation to the evenly spaced matrix of items which contained the target. One would expect inferences to be made later, at the semantic or conceptual level, rather than in the early stages of processing. It is therefore quite possible that visual attention operates both early and late in visual processing, and that both claims are partially correct.

The second area of debate, that the present functional approach to visual cognition might resolve, is related to the two main groups of visual cognition theories mentioned in the introduction to this thesis. These were the Gestalt theories and the Feature Analysis models. The Gestalt theories propose that objects are initially perceived as whole units and reduced into their component parts when necessary, while the Feature Analysis models propose that objects are initially encoded as separate features or elements which are subsequently conjoined to form the whole object we subjectively experience in visual perception. The present functional approach to visual cognition and feature integration proposes that visual space is divided into areas or units early in visual processing by perceptual grouping principles, and that feature integration is constrained by these divisions so that features within one area or unit are processed in parallel. It has already been described how this functional approach is compatible with the evidence that supports the Feature Analysis models like FIT, and it is also compatible with Hubel and Wiesel's (1962) evidence that individual cortical cells are activated by single features, like lines of a certain orientation and length.

The functional approach described here is also compatible with evidence that supports the Gestalt theories. For example, evidence to suggest people recognize configurations faster than their parts (Pomerantz et al., 1977) might be expected if cues to the boundaries between objects are processed first so that they constrain feature integration. Rhamachandran's (1990) evidence, demonstrating the subjective experience of a boundary led to the features contained within it to be captured into the apparent motion of that boundary, is also compatible with this functional approach to visual cognition as well as with the idea that configurations might be perceived faster than their parts. It is therefore suggested that both groups of theories have offered important hypotheses about visual cognition and, although they have appeared to have had very different views of how we perceive the world, the current evidence suggests these two approaches are reconcilable, each being right in some respects. This reconciliation is apparent in the suggested functional approach to visual cognition, that combines an extended version of Prinzmetal and Keysar's (1989) functional explanation of illusory conjunctions with Wolfe et al.'s (1989) Guided Search Model of feature integration.

SUMMARY

The results of two experiments were more readily explained by Prinzmetal and Keysar's functional explanation of illusory conjunctions than by FIT. That is, illusory conjunctions appear to arise as an adaptive response to very brief viewing times in the presence of constraining visual information which divides visual space into areas or units. When an extended version of Prinzmetal and Keysar' functional explanation (that includes the notion of local as well as global constraints) is combined with Wolfe et al.'s (1989) Guided Search Model of feature integration, the resulting functional approach to feature integration[2] can explain the pattern of non conjunction error rates as well as illusory conjunction rates across the four conditions utilized in the present experiments. It is suggested that constraints on visual processing have evolved from humans perceiving and interacting with the real world environment, and are therefore related to external factors that constitute cues to the boundaries between objects and areas of space. Also, some suggestions as to which environmental cues might constitute such constraints are made. It is further suggested that we may be hard-wired to divide peceptual information into units or chunks, and so may divide visual space into areas even when no obvious environmental cues indicate where the boundaries of areas might be. In such cases, top-down processes may be involved in constraining feature integration and attention. It is also

[2]NOTE: Since writing this thesis I have discovered that in 1973 Kahneman proposed a model of visual cognition and attention very similar, and probably more extensive, than the one described here. However, time has prevented me from investigating and discussing Kahneman's model in this thesis.

inferred from the present results, and a functional approach to visual cognition, that visual attentional selection occurs early in visual processing and is spatially based. However, this is not incompatable with the idea that it would be adaptive for organisms to attend to objects and further, that it is also possible that attentional selection might occur both early and late in visual processing. Finally, it is proposed that the functional approach to visual cognition discussed earlier, along with the present experimental results and Prinzmetal and Keysar's (1989) findings, suggests how the apparent conflict between the Gestalt theories and Feature Analysis models of visual cognition, might be resolved.

REFERENCES

Allport, A. (1989). Visual Attention. In M.I. Posner (Ed.), Foundations of Cognitive Science (p631-682). Cambridge, Massachusetts: The MIT Press.

Anderson, J.R. (1985). Cognitive Psychology and Its Implications. New York: Freeman.

Callaghan, T.C. (1989). Interference and dominance in texture segregation: Hue, geometric form and line orientation. Perception and Psychophysics, 46, 299-311.

Cave, K.R., & Wolfe, J.M. (1990). Modelling the role of parallel processing in visual search. Cognitive Psychology, 22, 225-271.

Cheal, M., & Lyon, D. (1989). Attention effects on form discrimination at different eccentricities. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41(4A), 719-746.

Cohen, A., & Ivry (1989). Illusory conjunctions inside and outside the focus of attention. Journal of Experimental Psychology: Human Perception and Performance, 15, 650-663.

Cohen, A., & Ivry, R. (1991). Density effects in conjunction search: Evidence for a course location mechanism in feature integration. Journal of Experimental Psychology: Human Perception and Performance, 17, 891-901.

Downing, C.J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. Journal of Experimental Psychology: Human Perception and Performance, 14, 188-202.

Downing, C.J., & Pinker, S. (1985). The spatial structure of visual attention: In M.I. Posner & O.S.M. Marin (Ed.), Mechanisms of Attention: Attention and Performance XI (p 171-187). Hillsdale, NJ: Erlbaum.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. Psychological Review, 87, 272-300.

Duncan, J. (1989). Boundary conditions on parallel processing in human vision. Perception, 18, 457-469.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. Psychological Review, 96, 433-458.

Egly, R., & Homa, D. (1991). Reallocation of visual attention. Journal of Experimental Psychology: Human Perception and Performance, 17, 142-159.

Enns, J.T. (1986). Seeing textons in context. Perception and Psychophysics, 39, 143-147.

Enns, J.T., & Rensink, R.A. (1990). Influence of scene-based properties on visual search. Science, 247, 721-723.

Erickson, C.W., & Yeh, Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance, 11, 583-597.

Flavell, J.H. (1985). Cognitive Development (2nd Edition). Englewood Cliffs, NJ: Prentice Hall.

Hubel, D.H., & Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 166, 106-154.

Humphreys, G.W., Riddock, M.J., & Quinlan, P.T. (1985). Interactive processes in perceptual organization: Evidence from visual agnosia. In M.I. Posner & O.S. Marin (Eds.), Attention and Performance XI. Hillsdale, NJ: Erlbaum.

Humphreys, G.W., & Bruce, V. (1989). Visual Cognition: Computational, experimental and neuropsychological perspectives. Hove: Erlbaum.

Johnston, W.A., Hawley, K.J., Plewe, S.H., Elliott, J.M.G., & DeWitt, M.J. (1990). Attention capture by novel stimuli. Journal of Experimental Psychology: General, 119, 397-411.

Johnston, J.C., & Pashler, H. (1990). Close binding of identity and location information in visual feature perception. Journal of Experimental Psychology: Human Perception and Performance, 16, 843-856.

Julesz, B. (1975). Experiments in the visual perception of texture. Scientific American, 232, 34-43.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. Nature, 290, 91-97.

Julesz, B. (1984). A brief outline of the Texton Theory of human vision. Trends in Neurosciences, 7, 41-45.

Julesz, B., & Bergen, J.R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. Bell System Technical Journal, 62, 1619-1645.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, N.J.: Prentice Hall.

LaBerge, D. (1983). Spatial extent of attention to letters and words. Journal of Experimental Psychology: Human Perception and Performance, 9, 371-379.

LaBerge, D., & Brown, V. (1986). Variations in size of the field in which targets are presented: An attentional range effect. Perception and Psychophysics, 40, 188-200.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. Psychological Review, 96, 101-124.

Lindsay, P.H., & Norman, D.A. (1972). Human Information Processing. New York: Academic Press.

Marr, D. (1982). Vision. San Francisco: Freeman.

McIlhagga, W., Hine, T., Cole, G.R., & Snyder, A.W. (1990). Texture segregation with luminance and chromatic contrast. Vision Research, 30, 489-495.

McLeod, P., Driver, J., Dienes, Z., & Crisp, J. (1991). Filtering by movement in visual search. Journal of Experimental Psychology, 17, 55-64.

Medin, D.L., & Wattenmaker, W.D. (1987). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), Concepts and Conceptual Development: Ecological and intellectual factors in categorization. Cambridge: Cambridge University Press.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Murphy, T.D., & Eriksen, C.W. (1987). Temporal changes in the distribution of attention in the visual field in response to precues. Perception and Psychophysics, 42, 576-586.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. Vision Research, 29, 1631-1647.

Nakayama, K., & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. Nature, 320, 264-265.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), Concepts and Conceptual Development: Ecological and intellectual factors in categorization. Cambridge: Cambridge University Press.

Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. Perception and Psychophysics, 41, 191-201.

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. Perception and Psychophysics, 43, 307-318.

Pomerantz, J.P., Sager, L.C., & Stoever, R.S. (1977). Perception of wholes and their component parts: Some configural superiority effects. Journal of Experimental Psychology: Human Perception and Performance, 3, 422-435.

Posner, M.I. (1980). Orienting of attention. The Quarterly Journal of Experimental Psychology, 32, 3-25.

Posner, M.I., Snyder, C.C.R., & Davidson, B.J. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. Perception and Psychophysics, 30, 330-340.

Prinzmetal, W., Hoffman, H., & Vest, K. (1991). Automatic processes in word perception: An analysis from illusory conjunctions. Journal of Experimental Psychology: Human Perception and Performance, 17, 902-923.

Prinzmetal, W., & Keysar, B. (1989). Functional theory of illusory conjunctions and neon colors. Journal of Experimental Psychology: General, 118, 165 190.

Prinzmetal, W., Presti, D.E., & Posner, M.I. (1986). Does attention affect feature integration? Journal of Experimental Psychology: Human Perception and Performance, 12, 361-369.

Prinzmetal, W., Treiman, R., & Rho, S.H. (1986). How to see a reading unit. Journal of Memory and Language, 25, 461-475.

Quinlan, P.T., & Humphreys, G.W. (1987). Visual search for targets defined by combinations of color, shape and size: An examination of the constraints on feature conjunction searches. Perception and Psychophysics, 41, 455-472.

Ramachandran, V.S. (1990). Visual perception in people and machines. In A. Blake & T. Troskianko (Eds.), AI and the Eye. Chichester: Wiley.

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Shulman, G.L., Remington, R.W., & McLean, J.P. (1979). Moving attention through visual space. Journal of Experimental Psychology, 5, 522-526.

Townsend, J.T. (1990). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. Psychological Science, 1, 46-54.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology: Human Perception and Performance, 8, 194-214.

Treisman, A. (1985). Properties, parts, and objects. In K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), Handbook of Perception and Human Performance, Vol. II: Cognitive Processes and Performance. New York: Wiley.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. Quarterly journal of Experimental Psychology, 40A, 201-237.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology, 12, 97-136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. Psychological Review, 95, 15-48.

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. Journal of Experimental Psychology: Human Perception and Performance, 10, 12-31.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance, 16, 459-478.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. Cognitive Psychology, 14, 107-141.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology: General, 114, 285-310.

Tsal, Y. (1983). Movements of attention across the visual field. Journal of Experimental Psychology: Human Perception and Performance, 9, 523-530.

Tsal, Y. (1989). Do illusory conjunctions support the feature integration theory? A critical review of theory and findings. Journal of Experimental Psychology: Human Perception and Performance, 15, 394-400.

Tsal, Y. (1989). Further comments on feature integration: A reply to Briand and Klein. Journal of Experimental Psychology: Human Perception and Performance, 15, 407-410.

Virzi, R.A., & Egeth, H.E. (1984). Is meaning implicated in illusory conjunctions? Journal of Experimental Psychology: Human Perception and Performance, 10, 573-580.

Wolfe, J.M., & Franzel, S.L. (1988). Binocularity and visual search. Perception and Psychophysics, 44, 81-93.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided Search: An alternative to the Feature Integration Model for visual search. Journal of Experimental Psychology: Human Perception and Performance, 15, 419-433.

Ziff, P. (1964). About ungrammaticalness. Mind, 73, 204-214.

APPENDICES

APPENDIX A: METHOD

1. Ages

Table 1: Each subject's age (in years), and mean age (rounded to nearest year) for Experiment 1 and Experiment 2.

Subject	Expt. 1 Age (in years)	Expt. 2 Age (in years)
1	20	37
2	32	17
3	43	20
4	37	24
5	24	36
6	24	43
7	23	19
8	21	20
9	26	20
10	24	19
11	27	19
12	25	20
13	43	32
14	24	28
15	23	23
16	29	24
17	23	20
18	22	24
19	36	23
20	21	19
21	23	20
22	19	20
23	17	23
24	18	19
Mean age	26	24

2.

Two preliminary test subjects were run through the experiment. It was found that the overhead fluorescent lighting in the experimental room was too bright for the slide projections, causing the images to appear faded. When the room was completely darkened (i.e. with no lighting and blackout curtains over the windows) the test subjects found it difficult to look at the screen for any length of time because of the brightness of the light from the projectors. This situation might also result in dark adaption which could have affected colour perception. The lamp used resulted in enough light to diffuse the bright light from the projections without being so bright as to fade the projected images.

3.

The slides used in these experiments were created by using ink-pens and stencils to draw two coloured letters on each of 144 A4 size sheets of matt finished cardboard. These were then photographed to produce the slides. The 144 cards containing the coloured letters then had the black square outlines added to them with a ruler and ink-pen. The cards were then photographed a second time to produce the slides for the three shape conditions.

4.

Visual angles were calculated by first calculating the distance subjects would need to be from the screen given the visual angle between the two letters required for each experiment. The formula used was $d = w/\tan^{-1} \emptyset$ where

d = the distance, in centimeters, from eyes to screen,

w = the distance, in centimeters, between the two letters, and

$\emptyset$ = the visual angle required for the distance between the letters.

The other visual angles were then calculated with the formula $\emptyset = \tan w/d$, where

$\emptyset$ = the visual angle,

w = the length, in centimeters, on the screen surface, and

d = the distance, in centimeters, between the subjects eyes and the screen.

5.

When the two preliminary subjects were run through the experiment it became apparent that a practice effect continued to affect their performance throughout the experimental blocks, so that they continued to improve to an extent where their total error rate could drop to a level that would not produce sufficient data. For one test subject, and some experimental subjects, this effect was quite large. Ideally subjects could have continued practice blocks until this practice effect reached a ceiling, but practical time considerations, particularly fatigue and peoples' willingness to participate in the experiment, prevented this.

APPENDIX B: RESULTS

1. Presentation times.

Table 1: Each subject's mean presentation time in 60ths. of a second, and the overall mean presentation time for both experiments.

Subject	Expt. 1. Mean pres. time (60ths sec.)	Expt. 2. Mean pres. time (60ths sec.)
1	2	3
2	6	2
3	4	6
4	4	4
5	5	5
6	6	6
7	5	2
8	6	2
9	5	2
10	4	3
11	4	2
12	5	4
13	6	6
14	4	5
15	6	6
16	3	6
17	5	3
18	5	2
19	6	4
20	4	3
21	5	4
22	4	3
23	5	3
24	3	5
Mean	4.67	3.79

2. t-tests.

One-tailed t-tests to test for significant levels above or below chance of illusory conjunctions, were performed using the formula D/SD/√n, where,

D = the mean of the differences between the colour conjunction errors and 0.5 of the colour feature errors,

SD = the standard deviation of the differences, and

√n = the square root of the sample size.

Table 1: Results of T-tests for significant levels of illusory conjunctions in each of the 4 conditions in both experiments (n = 24). An asterix marks the significant results.

Condition	Experiment 1.	Experiment 2.
No-shape	-.005/.029/√24 = -.847	-.002/.019/√24 = -.513
In-shape	.027/.044/√24 = 3.00 *	-.009/.028/√24 = -1.58
Out-shape	.046/.036/√24 = 6.21*	.006/.034/√24 = .87
Divided-shape	-.069/.033/√24 = -4.89*	-.036/.036/√24 = -5.14*

(See appendix B, 3, table 1 for a table of subjects' differences between illusory conjunction and 0.5 of colour feature error rates for Experiment 1, and appendix B, 3, table 3 for a table of these differences for Experiment 2.)

3. Differences between illusory conjunction & colour feature error rates.

TABLE 1: Differences in rates of illusory conjunctions and 0.5 of colour feature errors, as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.003	.00	.021	-.052
2	.01	-.021	.052	-.083
3	-.003	.063	.031	-.104
4	-.038	-.01	.042	-.031
5	.042	-.021	.052	-.042
6	-.021	.01	.094	-.125
7	-.028	-.01	.063	-.031
8	-.038	.094	.01	-.063
9	.003	-.021	-.021	-.083
10	-.028	.042	.083	.00
11	-.021	-.01	.00	-.063
12	.052	.052	.083	-.094
13	.01	.083	.083	-.073
14	-.028	.073	.083	-.083
15	-.017	.042	.021	-.094
16	.017	.094	.021	-.042
17	.028	.063	.00	-.01
18	-.007	-.042	.052	-.104
19	-.003	-.01	.063	-.073
20	.01	.052	.063	-.115
21	-.059	.021	.094	-.063
22	-.038	.104	.083	-.073
23	-.01	-.021	.063	-.042
24	.045	.01	-.021	-.115
Means	-.005	.026	.046	-.069
Stdevs.	.029	.044	.036	.033

TABLE 2: ANOVA summary table for the means of differences in rates of illusory conjunctions and 0.5 of colour feature errors, compared across 4 conditions in experiment 1 (n = 24)

Source of variation	df	Sum of squares	Mean square	F	p	Epsilon correction
Subjects	23	.025	.001			
Conditions	3	.183	.061	45.271	.0000	
Error	69	.093	.001			.97

TABLE 3: Results of Tukey tests for the means of the differences between illusory conjunctions and 0.5 of colour feature errors, compared across across 4 conditions in Experiment 1 (n = 24). Significant differences indicated are at the 0.05 level, or lower, where s indicates a significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	s	s	s
In-shape	s	X	-	s
Out-shape	s	-	X	s
Divided-shape	s	s	s	X

TABLE 1: Differences in rates between illusory conjunctions and 0.5 of colour feature errors, as proportions of each subject's total responses in 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	-.014	.021	.00	-.063
2	.017	.01	.01	-.01
3	-.045	.00	-.063	-.146
4	-.003	-.031	-.021	-.031
5	.014	-.052	.01	.00
6	-.007	-.063	.01	-.094
7	.024	-.01	-.021	-.01
8	-.007	-.021	.00	-.01
9	-.007	.021	.031	-.031
10	-.007	-.021	-.01	-.031
11	.014	.00	-.031	.00
12	.003	-.021	.031	.01
13	.01	.00	.115	-.094
14	-.007	.052	-.031	-.021
15	-.014	-.01	-.01	-.083
16	-.024	.00	.00	-.021
17	-.045	-.021	-.021	-.031
18	.003	-.063	.01	-.031
19	.00	.01	.031	-.042
20	.038	.021	.021	-.021
21	-.01	.01	.021	-.021
22	.003	-.042	.042	-.031
23	.003	-.01	.031	-.031
24	.007	.01	-.021	-.031
Means	-.002	-.009	.006	-.036
Stdevs.	.019	.028	.034	.036

TABLE 2: ANOVA summary table for the means of the differences in rates of illusory conjunctions and 0.5 of colour feature errors, compared across 4 conditions in Experiment 2 (n = 24)

Source of variation	df	Sum of squares	Mean square	*F*	*p*	Epsilon correction
Subjects	23	.027	.001			
Conditions	3	.024	.008	9.985	.0000	
Error	69	.056	.001			.79

TABLE 3: Results of Tukey tests for the means of the differences between colour conjunctions and 0.5 of colour feature errors, compared across 4 conditions in Experiment 2 (n = 24). Significant differences indicated are at the 0.01 level, where s indicates a significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	-	-	s
In-shape	-	X	-	s
Out-shape	-	-	X	s
Divided-shape	s	s	s	X

Note: Appendix B, 7 and 8, Tables 1 and 2, contain subjects' raw data for illusory conjunctions and colour feature errors.

4. Total error rates.

TABLE 1: Total error rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.285	.25	.333	.25
2	.424	.333	.313	.229
3	.306	.292	.229	.333
4	.243	.271	.167	.167
5	.243	.188	.104	.188
6	.375	.313	.354	.458
7	.208	.146	.25	.167
8	.361	.333	.229	.25
9	.299	.167	.333	.333
10	.299	.25	.146	.104
11	.285	.208	.188	.188
12	.375	.271	.271	.292
13	.361	.333	.354	.438
14	.34	.292	.229	.271
15	.278	.229	.229	.229
16	.34	.313	.333	.354
17	.396	.375	.333	.313
18	.444	.25	.354	.375
19	.257	.25	.292	.25
20	.354	.271	.271	.375
21	.375	.417	.354	.313
22	.313	.188	.271	.229
23	.208	.167	.188	.167
24	.368	.25	.333	.395
Means	.322	.265	.269	.278
Stdevs.	.063	.067	.072	.091

TABLE 2: ANOVA summary table for mean total error rates compared across 4 conditions in Experiment 1 (n = 24).

Source of variation	df	Sum of squares	Mean square	F	p	Epsilon correction
Subjects	23	.363	.016			
Conditions	3	.050	.017	7.189	.0003	
Error	69	.161	.002			.74

TABLE 3: Tukey test results for mean total rates of errors compared across 4 conditions in Experiment 1 (n = 24). Significant differences indicated are at the 0.05 level or lower, where s indicates a significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	s	s	s
In-shape	s	X	-	-
Out-shape	s	-	X	-
Divided-shape	s	-	-	X

TABLE 4: Total error rates as a proportion of each subject's total responses in each of 4 conditions, as well as means and standard deviations, for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.243	.188	.208	.292
2	.243	.146	.104	.271
3	.208	.146	.208	.333
4	.181	.146	.063	.146
5	.125	.146	.083	.125
6	.16	.125	.208	.188
7	.111	.125	.125	.125
8	.153	.063	.083	.125
9	.16	.104	.083	.083
10	.188	.125	.042	.104
11	.111	.146	.104	.083
12	.125	.083	.146	.146
13	.368	.417	.333	.438
14	.16	.125	.167	.188
15	.319	.25	.333	.4
16	.243	.125	.167	.271
17	.188	.146	.104	.167
18	.222	.146	.146	.104
19	.188	.083	.167	.146
20	.215	.229	.167	.188
21	.188	.167	.083	.146
22	.181	.146	.167	.208
23	.181	.104	.125	.104
24	.236	.104	.146	.208
Means	.196	.149	.148	.191
Stdevs.	.061	.071	.073	.096

TABLE 5: ANOVA summary table for a comparison of mean total error rates across 4 conditions in Experiment 2 (n = 24).

Source of variation	df	Sum of squares	Mean square	F	p	Epsilon correction
Subjects	23	.429	.019			
Conditions	3	.048	.016	10.037	.0000	
Error	69	.109	.002			.86

TABLE 6: Tukey test results of a comparison of mean total error rates across 4 conditions in Experiment 2 (n = 24). Significant results are at the 0.05 level or less, where s indicates a significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	s	s	-
In-shape	s	X	-	s
Out-shape	s	-	X	s
Divided-shape	-	s	s	X

5. Feature and complete errors (Non conjunction errors).

TABLE 1: Non conjunction errors, as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n =24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.229	.187	.25	.25
2	.334	.25	.209	.229
3	.25	.188	.146	.333
4	.222	.208	.104	.167
5	.153	.146	.041	.188
6	.312	.23	.187	.458
7	.201	.125	.167	.167
8	.305	.187	.166	.25
9	.173	.125	.27	.312
10	.257	.167	.042	.083
11	.243	.166	.146	.188
12	.25	.167	.146	.292
13	.178	.187	.229	.396
14	.277	.167	.104	.25
15	.222	.146	.166	.229
16	.264	.167	.25	.312
17	.292	.25	.27	.25
18	.368	.208	.25	.375
19	.215	.208	.188	.25
20	.271	.167	.146	.375
21	.306	.271	.187	.271
22	.264	.063	.146	.208
23	.159	.125	.084	.167
24	.264	.167	.27	.395
Means	.25	.178	.174	.266
Stdevs.	.055	.046	.068	.09

TABLE 2: ANOVA summary table for the comparison of the mean non conjunction error rates, across 4 conditions in Experiment 1 (n = 24).

Source of variation	df	Sum of squares	Mean square	F	p	Epsilon correction
Subjects	23	.216	.009			
Conditions	3	.167	.056	19.745	.0000	
Error	69	.195	.003			.75

Table 3: Tukey test results for a comparison of mean non conjunction error rates, across 4 conditions in Experiment 1 (n = 24). Significant differences are at the 0.01 level, where s indicates a significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	s	s	-
In-shape	-	X	-	s
Out-shape	s	-	X	s
Divided-shape	-	s	s	X

TABLE 4: Non conjunction error rates, as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.208	.125	.166	.292
2	.208	.104	.083	.229
3	.173	.104	.187	.333
4	.146	.125	.063	.146
5	.097	.146	.062	.104
6	.132	.125	.145	.188
7	.076	.104	.104	.125
8	.139	.063	.083	.125
9	.132	.062	.041	.083
10	.163	.125	.042	.104
11	.09	.125	.104	.083
12	.104	.083	.104	.104
13	.271	.313	.145	.396
14	.132	.062	.146	.167
15	.263	.208	.27	.4
16	.222	.104	.146	.229
17	.188	.146	.104	.167
18	.173	.146	.104	.104
19	.153	.062	.104	.146
20	.146	.166	.125	.188
21	.16	.125	.062	.146
22	.153	.146	.125	.208
23	.153	.083	.083	.104
24	.201	.083	.146	.208
Means	.162	.122	.114	.182
Stdevs.	.05	.054	.051	.092

TABLE 5: ANOVA summary table for a comparison of mean rates of non conjunction errors, across 4 conditions in Experiment 2 (n = 24).

Source of variation	df	Sum of squares	Mean square	F	p	Epsilon correction
Subjects	23	.259	.011			
Conditions	3	.075	.025	14.395	.0000	
Error	69	.120	.002			.85

TABLE 6: Tukey test results for a comparison of mean non conjunction error rates, across 4 conditions in Experiment 2 (n = 24). Significant differences indicated are at the 0.01 level, where s indicates significant difference.

	No-shape	In-shape	Out-shape	Divided-shape
No-shape	X	s	s	-
In-shape	s	X	-	s
Out-shape	s	-	X	s
Divided-shape	-	s	s	X

6. Raw data tables, Experiment 1.

Table 1: Colour feature error rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations of each condition in Experiment 1 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.104	.125	.125	.104
2	.16	.208	.104	.167
3	.118	.083	.104	.208
4	.118	.146	.042	.063
5	.097	.125	.021	.083
6	.167	.146	.146	.25
7	.069	.063	.042	.063
8	.188	.104	.104	.125
9	.104	.125	.167	.208
10	.139	.083	.042	.042
11	.125	.104	.083	.125
12	.146	.104	.083	.188
13	.146	.125	.083	.229
14	.181	.104	.083	.208
15	.146	.083	.083	.188
16	.118	.104	.125	.167
17	.153	.125	.125	.146
18	.167	.167	.104	.208
19	.09	.104	.083	.146
20	.146	.104	.125	.229
21	.257	.25	.146	.208
22	.174	.042	.083	.188
23	.118	.125	.083	.083
24	.118	.146	.167	.229
Means	.139	.121	.098	.161
Stdevs.	.039	.044	.039	.062

TABLE 2: Illusory conjunction rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n = 24).

	Conditions			
Subjects	No-shape	In-shape	Out-shape	Divided-shape
1	.056	.063	.083	0
2	.09	.083	.104	0
3	.056	.104	.083	0
4	.021	.063	.063	0
5	.09	.042	.063	0
6	.063	.083	.167	0
7	.007	.021	.083	0
8	.056	.146	.063	0
9	.056	.042	.063	.021
10	.042	.083	.104	.021
11	.042	.042	.042	0
12	.125	.104	.125	0
13	.083	.146	.125	.042
14	.063	.125	.125	.021
15	.056	.083	.063	0
16	.076	.146	.083	.042
17	.104	.125	.063	.063
18	.076	.042	.104	0
19	.042	.042	.104	0
20	.083	.104	.125	0
21	.069	.146	.167	.042
22	.049	.125	.125	.021
23	.049	.042	.104	0
24	.104	.083	.063	0
Means	.065	.087	.095	.011
Stdevs.	.027	.04	.034	.018

TABLE 3: Letter feature error rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n =24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.125	.042	.125	.146
2	.16	.042	.104	.063
3	.118	.104	.042	.125
4	.097	.063	.063	.104
5	.056	.021	.021	.104
6	.146	.083	.042	.188
7	.125	.063	.104	.104
8	.097	.083	.063	.125
9	.132	0	.104	.104
10	.097	.063	0	.042
11	.104	.063	.063	.063
12	.09	.042	.042	.083
13	.111	.042	.125	.146
14	.069	.063	.021	.021
15	.049	.042	.063	.021
16	.118	.042	.104	.104
17	.132	.125	.146	.083
18	.146	.042	.125	.146
19	.097	.083	.083	.104
20	.111	.063	.021	.125
21	.049	.021	.042	.063
22	.076	.021	.063	.021
23	.028	0	0	.083
24	.118	0	.104	.146
Means	.102	.05	.069	.096
Stdevs.	.034	.032	.042	.044

TABLE 4: Complete error rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	0	.021	0	0
2	.014	0	0	0
3	.007	0	0	0
4	.007	0	0	0
5	0	0	0	0
6	0	0	0	.021
7	.007	0	.021	0
8	.021	0	0	0
9	0	0	0	0
10	.021	.021	0	0
11	.007	0	0	0
12	.007	.021	0	.021
13	.014	.021	.021	.021
14	.021	0	0	.021
15	.028	.021	.021	.021
16	.028	.021	.021	.021
17	.007	0	0	.021
18	.042	0	0	0
19	.028	0	.021	0
20	.007	0	0	.021
21	0	0	0	0
22	.007	0	0	0
23	.007	0	0	0
24	.014	0	0	.021
Means	.012	.005	.004	.008
Stdevs.	.011	.009	.009	.01

TABLE 5: Letter feature with colour conjunction error rates, as a proportion of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 1 (n = 24).

	Conditions			
Subjects	No-shape	In-shape	Out-shape	Divided-shape
1	0	0	0	0
2	0	0	0	0
3	.007	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	.007	0	0	0
10	0	0	0	0
11	.007	0	0	0
12	.007	0	.021	0
13	.007	0	0	0
14	.007	0	0	0
15	0	0	0	0
16	0	0	0	.021
17	0	0	0	0
18	.014	0	.021	.021
19	0	.021	0	0
20	.007	0	0	0
21	0	0	0	0
22	.007	0	0	0
23	.007	0	0	0
24	.014	.021	0	0
Means	.004	.002	.002	.002
Stdevs.	.005	.006	.006	.006

7. Raw data tables, Experiment 2.

TABLE 1: Colour feature error rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subjects	No-shape	In-shape	Out-shape	Divided-shape
1	.097	.083	.083	.125
2	.035	.063	.021	.104
3	.16	.083	.167	.292
4	.076	.104	.042	.063
5	.028	.104	.021	.042
6	.069	.125	.104	.188
7	.021	.063	.083	.021
8	.042	.042	0	.021
9	.069	.042	.021	.063
10	.056	.042	.021	.063
11	.012	.042	.063	0
12	.035	.042	.021	.063
13	.174	.021	.146	.271
14	.069	.021	.104	.083
15	.139	.104	.146	.167
16	.09	.042	.042	.125
17	.09	.042	.042	.063
18	.09	.125	.063	.063
19	.069	.021	.063	.083
20	.063	.083	.042	.042
21	.076	.063	0	.042
22	.049	.083	0	.063
23	.049	.063	.021	.063
24	.056	.021	.042	.063
Means	.071	.071	.056	.09
Stdevs.	.04	.043	.048	.073

TABLE 2: Illusory conjunction rates as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subjects	No-shape	In-shape	Out-shape	Divided-shape
1	.035	.063	.042	0
2	.035	.042	.021	.042
3	.035	.042	.021	0
4	.035	.021	0	0
5	.028	0	.021	.021
6	.028	0	.063	0
7	.035	.021	.021	0
8	.014	0	0	0
9	.028	.042	.042	0
10	.021	0	0	0
11	.021	.021	0	0
12	.021	0	.042	.042
13	.097	.104	.188	.042
14	.028	.063	.021	.021
15	.056	.042	.063	0
16	.021	.021	.021	.042
17	0	0	0	0
18	.049	0	.042	0
19	.035	.021	.063	0
20	.069	.063	.042	0
21	.028	.042	.021	0
22	.028	0	.042	0
23	.028	.021	.042	0
24	.035	.021	0	0
Means	.034	.027	.034	.009
Stdevs.	.019	.027	.039	.016

TABLE 3: Letter feature error rates, as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.083	.021	.042	.104
2	.111	.021	.063	.104
3	.014	.021	.021	.042
4	.069	.021	.021	.083
5	.069	.042	0	.063
6	.021	0	.042	0
7	.056	.042	.021	.104
8	.083	.021	.063	.104
9	.049	.021	.021	0
10	.069	.042	0	.042
11	.069	.083	.042	.083
12	.069	.042	.083	.021
13	.063	.063	0	.104
14	.063	.042	.042	.083
15	.104	.083	.104	.188
16	.111	.063	.083	.104
17	.069	.083	.042	.063
18	.069	.021	.042	.042
19	.076	.042	.042	.042
20	.083	.083	.083	.104
21	.069	.042	.063	.083
22	.097	.063	.125	.146
23	.076	.021	.063	.042
24	.139	.063	.104	.146
Means	.074	.043	.05	.079
Stdevs.	.027	.025	.034	.046

TABLE 4: Complete error rates, as proportions of each subject's total responses in each of 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.021	.021	.042	.042
2	.042	.021	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	.042	0
6	.028	0	0	0
7	0	0	0	0
8	.014	0	.021	0
9	.014	0	0	.021
10	.042	.042	.021	0
11	.007	0	0	0
12	0	0	0	0
13	.028	0	0	.021
14	0	0	0	0
15	.021	.021	.021	.021
16	.014	0	.021	0
17	.028	.021	.021	.042
18	.014	0	0	0
19	0	0	0	0
20	0	0	0	.021
21	.007	0	0	0
22	.007	0	0	0
23	.021	0	0	0
24	.007	0	0	0
Means	.014	.005	.008	.007
Stdevs	.013	.011	.013	.013

TABLE 5: Letter feature with colour conjunction error rates, as proportions of each subject's total responses in 4 conditions, as well as means and standard deviations for each condition in Experiment 2 (n = 24).

	Conditions			
Subject	No-shape	In-shape	Out-shape	Divided-shape
1	.007	0	0	.021
2	.021	0	0	.021
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	.014	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	0	0
10	0	0	0	0
11	0	0	0	0
12	0	0	0	.021
13	.007	.042	0	0
14	0	0	0	0
15	0	0	0	.021
16	.007	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	.007	0	0	.021
21	.007	.021	0	.021
22	0	0	0	0
23	.007	0	0	0
24	0	0	0	0
Means	.003	.003	0	.006
Stdevs.	.005	.009	0	.01